the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2021) doi: 10.1111/tpj.15292

The bundle sheath of rice is conditioned to play an active role in water transport as well as sulfur assimilation and jasmonic acid synthesis

Lei Hua[1], Sean R. Stevenson[1], Ivan Reyna-Llorens[1], Haiyan Xiong[1], Stanislav Kopriva[2] and Julian M. Hibberd[1,*]

[1]*Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EA, UK, and*

[2]*Institute for Plant Sciences, Cluster of Excellence on Plant Sciences (CEPLAS), University of Cologne, Zülpicher Str. 47b, Cologne 50674, Germany*

Received 5 March 2021; revised 16 April 2021; accepted 20 April 2021.

*For correspondence (e-mail jmh65@cam.ac.uk).

SUMMARY

Leaves comprise multiple cell types but our knowledge of the patterns of gene expression that underpin their functional specialization is fragmentary. Our understanding and ability to undertake the rational redesign of these cells is therefore limited. We aimed to identify genes associated with the incompletely understood bundle sheath of C_3 plants, which represents a key target associated with engineering traits such as C_4 photosynthesis into *Oryza sativa* (rice). To better understand the veins, bundle sheath and mesophyll cells of rice, we used laser capture microdissection followed by deep sequencing. Gene expression of the mesophyll is conditioned to allow coenzyme metabolism and redox homeostasis, as well as photosynthesis. In contrast, the bundle sheath is specialized in water transport, sulphur assimilation and jasmonic acid biosynthesis. Despite the small chloroplast compartment of bundle sheath cells, substantial photosynthesis gene expression was detected. These patterns of gene expression were not associated with the presence or absence of specific transcription factors in each cell type, but were instead associated with gradients in expression across the leaf. Comparative analysis with C_3 Arabidopsis identified a small gene set preferentially expressed in the bundle sheath cells of both species. This gene set included genes encoding transcription factors from 14 orthogroups and proteins allowing water transport, sulphate assimilation and jasmonic acid synthesis. The most parsimonious explanation for our findings is that bundle sheath cells from the last common ancestor of rice and Arabidopsis were specialized in this manner, and as the species diverged these patterns of gene expression have been maintained.

Keywords: *Oryza sativa*, bundle sheath, water transport, sulphur metabolism, deep sequencing, laser capture microdissection.

INTRODUCTION

Although the major cell types of leaves were described in the 19th century (Haberlandt, 1884) we have an incomplete understanding of the role that each plays (Aubry *et al.*, 2014b; Mustroph *et al.*, 2009). This lack of knowledge hinders our understanding of how basic processes are organized but is also likely to limit the rational redesign of leaves for crop improvement. One example is associated with attempts to engineer C_4 photosynthesis into species such as C_3 *Oryza sativa* (rice) to increase yields (von Caemmerer *et al.*, 2012; Hibberd *et al.*, 2008; Langdale, 2011). As C_4 photosynthesis typically requires metabolic compartmentation between mesophyll and bundle sheath cells, a better understanding of these tissues in rice may facilitate such a project.

The C_4 pathway is thought to have evolved over 60 times independently in monocotyledons and dicotyledons, in response to reduced water availability and CO_2 supply (Sage, 2004). These environmental factors reduce the ratio of carboxylation to oxygenation events at the active site of RuBisCO, and so lead to higher rates of the photorespiration (Berry, 2012; Sage *et al.*, 2012; Tipple and Pagani, 2007). Although in some C_4 species a carbon-concentrating mechanism is established within large cells (von Caemmerer *et al.*, 2014; Jurić *et al.*, 2016; Voznesenskaya *et al.*, 2001), in the majority of known lineages this takes place across distinct cell types. In these two-celled C_4 species, RuBisCO activity is replaced with phosphoenolpyruvate carboxylase in mesophyll cells to allow bicarbonate to be fixed into oxaloacetate. High concentrations of C_4 acids

© 2021 The Authors.
The Plant Journal published by Society for Experimental Biology and John Wiley & Sons Ltd
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

derived from oxaloacetate build up in mesophyll cells and drive diffusion to the bundle sheath, where C_4 acid decarboxylases resupply CO_2 to RuBisCO. This reorganization of photosynthesis is thus enabled by the presence of distinct cell types such as the mesophyll and bundle sheath (Furbank, 2016). Not only do bundle sheath cells of C_4 plants undertake a key role in photosynthesis, but they are also the primary location of starch synthesis (Lunn and Furbank, 1997) and sulphur metabolism (Burgener *et al.*, 1998; Burnell, 1984; Gerwick *et al.*, 1980; Passera and Ghisi, 1982; Schmutz and Brunold, 1984). Distinct classes of bundle sheath cells have been reported in *Zea mays* (maize), with abaxial bundle sheath cells of rank-2 intermediate veins being specialized for phloem loading (Bezrutczyk *et al.*, 2021). The importance of bundle sheath cells in the C_4 leaf, and the discovery that their thickened cell walls allow them to be separated from the rest of the leaf, led to them being studied intensively. In summary, in the C_4 leaf the bundle sheath cells are well characterized and carry out a variety of specialized roles.

Although the vast majority of plants use the ancestral C_3 pathway, our understanding of gene expression in bundle sheath cells specifically, and other major cell types more generally, of C_3 leaves is poor. In *Arabidopsis thaliana* (hereafter Arabidopsis) the bundle sheath represents approximately 15% of chloroplast-containing cells in the leaf (Kinsman and Pyke, 1998). Although the photosynthetic apparatus is functional in C_3 bundle sheath cells (Fryer *et al.*, 2002; Williams *et al.*, 1989), the absolute number of chloroplasts per cell is low. Consistent with this, reducing chlorophyll accumulation or increasing the chloroplast compartment in these cells have limited impact on leaf-level photosynthesis (Janacek *et al.*, 2009; Wang *et al.*, 2017). Instead, it appears that the bundle sheath of Arabidopsis is specialized in sulphur metabolism and glucosinolate synthesis (Aubry *et al.*, 2014b; Koroleva *et al.*, 2010). The stress-responsive regulation of aquaporins in bundle sheath cells is considered important for the hydraulic conductance of the whole leaf (Attia *et al.*, 2020; Sade *et al.*, 2014; Shatil-Cohen *et al.*, 2011), and consistent with this the bundle sheath cells have been proposed to help maintain the hydraulic integrity of the xylem (Griffiths *et al.*, 2013; Sage, 2001) as well as regulate the flux of mineral and metabolites in and out of the leaf (Leegood, 2008; Wigoda *et al.*, 2017). In summary, compared with C_4 plants we have a relatively poor understanding of the role of bundle sheath cells in C_3 species, and this is particularly the case in grasses such as rice.

In roots, one approach that has been used widely to define the function of discrete cell types is to generate lines in which distinct tissues are marked with a fluorescent protein, and after protoplast isolation and cell sorting the patterns of gene expression are defined (Birnbaum *et al.*, 2003; Brady *et al.*, 2007). In leaves, this approach has been less widely adopted, probably because of the longer incubation times typically required to generate protoplasts and because of concerns about stress and de-differentiation taking place during protoplasting (Sawers *et al.*, 2007). Recently, rapid protoplasting followed by the isolation of bundle sheath cells indicated a key role in transport (Wigoda *et al.*, 2017), and single-cell RNA sequencing allowed distinct patterns of gene expression to be related to discrete cell types of the Arabidopsis leaf, indicating that bundle sheath protoplasts were most similar to xylem cells (Kim *et al.*, 2021).

Alternative technologies that have been applied to this problem include the isolation of ribosomes from specific cell types after they were labelled with an exogenous tag (Aubry *et al.*, 2014b; Mustroph *et al.*, 2009) or the use of laser capture microdissection (LCM) (Jiao *et al.*, 2009). This latter approach circumvents the need to identify promoters with highly specific expression domains and the production of transgenic lines.

Here we used an optimized LCM protocol for RNA isolation from leaves (Hua and Hibberd, 2019) to study the patterns of gene expression in bundle sheath, veinal (including phloem and xylem parenchyma, xylem, as well as sieve elements and companion cells) and mesophyll cells of rice. We had three main hypotheses. First, that as the rice bundle sheath contains a small chloroplast compartment, gene expression would be poorly set up for photosynthesis. Second, as in other species, gene expression in the rice bundle sheath would favour water transport. Third, in contrast to Arabidopsis, as rice does not make glucosinolates, there has been no selection pressure to restrict sulphur assimilation to the bundle sheath. Mesophyll and bundle sheath cells were distinguished by the over-representation of terms, including photosynthesis and co-enzyme metabolism in the mesophyll cells and solute transport and protein synthesis in the bundle sheath cells. Transcripts encoding the majority of aquaporins were more abundant in bundle sheath cells, and this was also the case for transcripts encoding proteins associated with sulphur assimilation and nitrate reduction. Transcription factors that were preferential to each cell type were identified, but in most cases a gradient from veins to bundle sheath to mesophyll cells, or *vice versa*, was detected. Direct comparison with publicly available data from Arabidopsis identified groups of genes encoding aquaporins, proteins allowing sulphur assimilation and jasmonic acid synthesis, and also a small number of transcription factor families that showed preferential expression in the bundle sheaths of both species. Although these findings could be linked to evolutionary convergence, the most parsimonious explanation is that the bundle sheath of the last common ancestor of monocotyledons and dicotyledons was specialized in water transport, sulphur assimilation and jasmonic acid synthesis, and that members of the basic

© 2021 The Authors.
The Plant Journal published by Society for Experimental Biology and John Wiley & Sons Ltd,

leucine zipper, C_2H_2-type zinc finger, DNA-binding with one finger, ethylene-responsive factor, hairy-related transcription factor, GRAS, MYB, nuclear factor YB and vascular plant one zinc finger protein transcription factor families play ancient and conserved roles in this cell type.

RESULTS

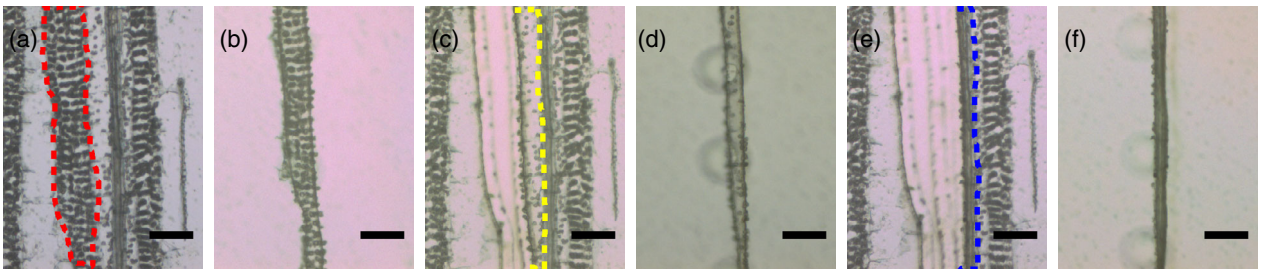
The rice bundle sheath is specialized for transport but also for photosynthesis

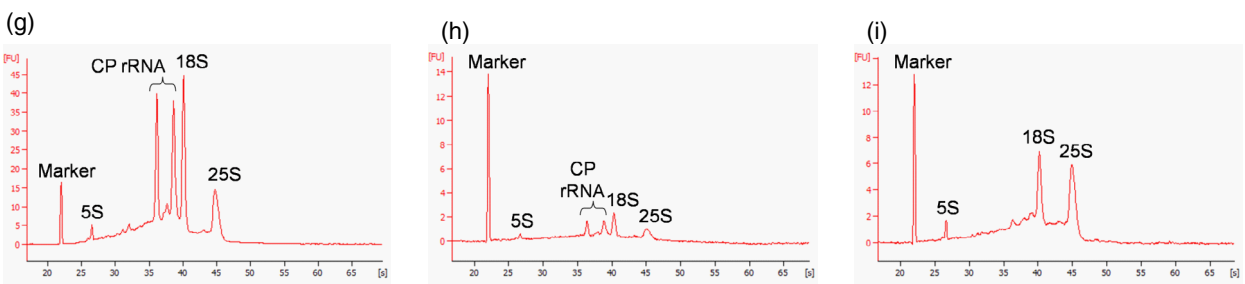
To gain insight into the genetic basis for functional specialization associated with mesophyll, bundle sheath and veinal cells of rice, LCM was used to isolate RNA from these tissues. Paradermal sections allowed the unambiguous identification of mesophyll cells (Figure 1a), bundle sheath cells containing large vacuoles and fewer chloroplasts (Figure 1c) and veins (Figure 1e). To isolate each cell type with minimal cross-contamination, mesophyll cells were first dissected and captured (Figure 1a,b). The sequential capture of bundle sheath cells (Figure 1c,d) followed by veinal cells was then possible (Figure 1e,f). RNA was extracted from each tissue and RNA integrity numbers ranging from 6.0 to 7.3 indicated good quality. Strong peaks associated with ribosomal RNAs of the chloroplast were evident in mesophyll cells (Figure 1g), whereas in bundle sheath cells they were less abundant (Figure 1h) and in veins they were not discernible (Figure 1i).

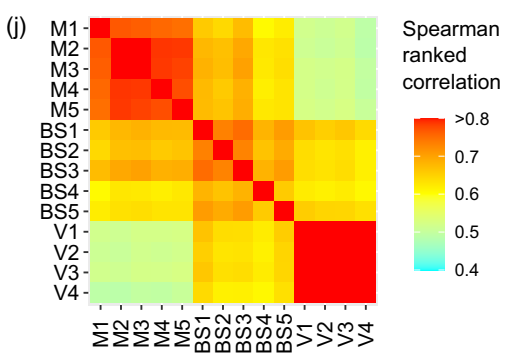
3′ mRNA sequencing was performed and from each cell type, 24–36 million reads from four or five biological replicates were obtained. After processing to remove low-quality reads, 13–23 million were quantified against the rice cDNA reference (MSU 7) (Table S1). An average of 10 097, 10 083 and 13 648 transcripts were detected in each cell type (Table S1). Spearman ranked correlation coefficients for gene expression showed little variation between replicates from each cell type, and each cell type exhibited distinct patterns of gene expression (Figure 1j). Principal component analysis (PCA) also showed the close grouping of biological replicates, and that 46.1% of variance was associated with the three cell types, whereas a second component separated the bundle sheath cells from the mesophyll and vein cells (Figure 1k). To assess the purity of the tissues sampled, we examined the transcript abundance of genes previously reported to be associated with each cell type. Consistent with these studies, sucrose phosphate synthase (*SPS1*, LOC_Os01g69030) and aquaporin *PIP2;7* (LOC_Os09g36930) were preferentially expressed in mesophyll cells (Chávez-Bárcenas *et al.*, 2000; Li *et al.*, 2008), transcripts derived from two tonoplast monosaccharide transporters *TMT1* and *TMT2* were detected in mesophyll, bundle sheath cells and vascular bundles (Cho *et al.*, 2010), phosphoenolpyruvate carboxykinase (*PCK1*) was expressed most strongly in the bundle sheath and veins (Nomura *et al.*, 2005), and the sucrose transporter *SUT1* (LOC_Os03g07480) (Ibraheem *et al.*, 2013; Scofield *et al.*, 2007) was predominately expressed in vascular tissue (Figure S1). The strong clustering of each cell type as well as congruence with previous studies are consistent with the notion that these samples obtained by LCM contained relatively little cross-contamination.

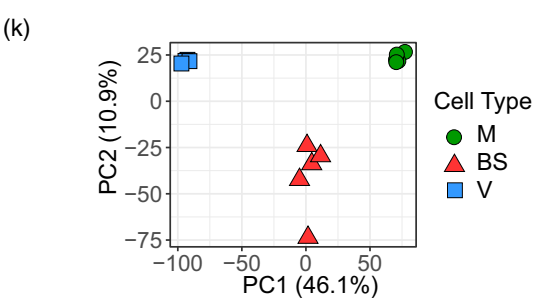
To quantify the extent to which transcript abundance differed between the bundle sheath, mesophyll and veinal cells, we performed differential gene expression analysis using DESEQ2 and EDGER (Love *et al.*, 2014; Robinson *et al.*, 2010). This identified 1919 differentially expressed genes between bundle sheath and mesophyll cells, the majority of which (1173) were more abundant in bundle sheath cells (false discovery rate and adjusted $P < 5\%$) (Tables S2 and S3). Functional enrichment analysis identified three categories over-represented in mesophyll cells containing transcripts linked to photosynthesis, coenzyme metabolism and redox homeostasis (Figure 1l). Eight categories were associated with bundle sheath cells, including solute transport, protein biosynthesis, phytohormone action, nutrient uptake, polyamine metabolism and cellular respiration (Figure 1l). To provide an overview of metabolic specialization in mesophyll and bundle sheath cells, MAPMAN (Schwacke *et al.*, 2019; Thimm *et al.*, 2004) was used. Consistent with mesophyll cells being specialized for photosynthesis, this indicated that transcripts encoding components of the light reactions, Calvin–Benson–Bassham cycle and tetrapyrolle biosynthesis were upregulated in mesophyll cells, whereas those associated with many other metabolic processes, including cell wall, minor carbohydrate, fatty acid, amino acid and secondary metabolism, were more abundant in the bundle sheath (Figure 1m). Quantitative comparison of bundle sheath and veinal cells showed 1258 and 660 genes were significantly up- and downregulated in bundle sheath cells, respectively (Tables S2 and S3), and indicated categories associated with the bundle sheath included photosynthesis, carbohydrate metabolism, redox homeostasis, secondary metabolism and solute transport (Figure S2a). MAPMAN outputs confirmed transcripts encoding components of the light-dependent reactions of photosynthesis, as well as the Calvin–Benson–Bassham cycle and photorespiratory pathway were more abundant in bundle sheath than in veinal cells (Figure S2b). In contrast, transcripts preferential to veins were associated with processes that included RNA biosynthesis, protein homeostasis, lipid metabolism and solute transport (Figure S2a). When mesophyll and veins were compared, a greater number of differentially expressed genes were identified with 1728 and 2038 transcripts being more abundant in mesophyll and veins, respectively (Tables S2 and S3). The expected preferential expression of photosynthesis-related genes in mesophyll cells was detected, and transcripts associated with protein biosynthesis and cellular respiration were

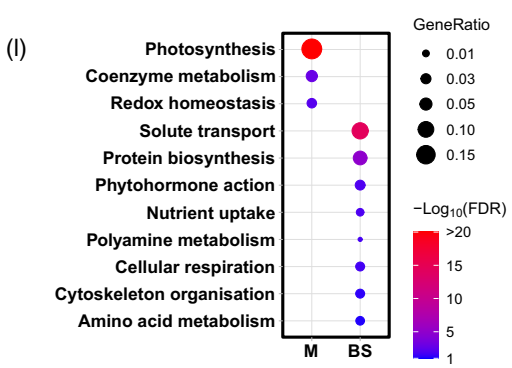
© 2021 The Authors.
The Plant Journal published by Society for Experimental Biology and John Wiley & Sons Ltd,

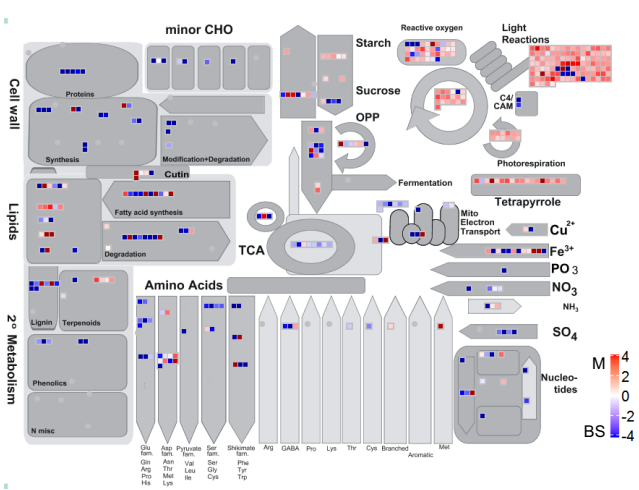


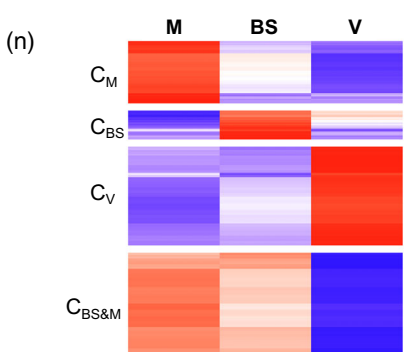


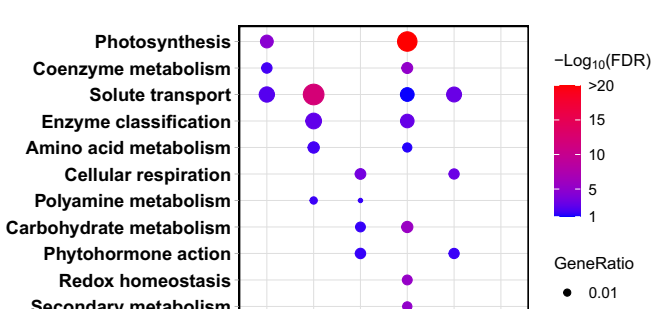












© 2021 The Authors.

Figure 1. RNA was isolated from mesophyll, bundle sheath and veinal cells of *Oryza sativa* (rice) using laser capture microdissection. Each cell type was identified and then sequentially removed from paradermal sections prior to RNA quality being assessed. (a, b) Representative image of mesophyll cells, outlined with a red dashed line (a) that are cut with an ultraviolet laser and then captured with an infrared laser and placed on a cap (b). (c, d) Bundle sheath cells (c) were then cut and captured (d). (e, f) Lastly, veinal cells (e) were cut and captured (f). Scale bars: 50 μm. (g–i) Representative RNA profiles from microdissected mesophyll (g), bundle sheath (h) and veinal cells (i). Peaks from the cytosolic *25S*, *18S* and *5S* ribosomal RNAs were detected in all cell types. Chloroplastic ribosomal RNAs (CP rRNAs) were clearly detectable in mesophyll and bundle sheath cells. (j) Spearman ranked correlations of $\log_2$-transformed transcripts per million (TPM) values indicate little variation between biological replicates from each cell type and distinct patterns of gene expression in each cell type. (k) Principal component analysis of normalized counts after variance-stabilizing transformation shows that cell type accounted for 46.1% of the variance detected in the data. (l) Primary MAPMAN categories associated with differentially expressed genes in bundle sheath and mesophyll cells. Terms were defined using Fisher's exact test (false discovery rate, FDR < 0.1), the colour scale represents negative $\log_{10}$-transformed FDR and the gene ratio represents the ratio of matched genes in categories relative to the total number of differentially expressed genes in each cell type. (m) Metabolic overviews of differentially expressed genes between bundle sheath and mesophyll cells. Colour scale presents the $\log_2$ fold change. (n) *K*-mean clustering of 4155 differentially expressed transcripts was performed using $\log_2$-transformed quantile-normalized transcripts per million (TPM) values and visualized in a heatmap. Clusters were named as C_M, C_{BS}, C_V, $C_{BS\&M}$, $C_{BS\&V}$ and $C_{M\&V}$. C_M contained 613 genes that were strongest in mesophyll cells, C_{BS} contained 285 genes with preferential expression in bundle sheath cells, C_V contained 972 genes that were most strongly expressed in veins, $C_{BS\&M}$ contained 1136 genes mostly highly expressed in mesophyll and bundle sheath cells, $C_{BS\&V}$ contained 1015 genes that were preferential to bundle sheath and veinal cells, and $C_{M\&V}$ contained 134 genes strongly expressed in both mesophyll and veinal cells. Colour scale represents *Z*-score. (o) Schematic illustrating the enriched primary categories derived from MAPMAN using Fisher's exact test (FDR < 0.1) for each of the six clusters: colour scale indicates negative $\log_{10}$-transformed FDR; size of dots (GeneRatio) represents the ratio of matched genes in each category relative to the total number of genes in each cluster.

more abundant in veins (Figure S2c,d). The greater number of differentially expressed genes between mesophyll and veins is consistent with the correlation and PCA analysis (Figure 1j,k).

To further assess patterns of transcript abundance across all three cell types, we clustered genes based on expression. A total of 4155 genes defined as being differentially expressed in the pairwise comparisons above were partitioned into six clusters associated with the cell types in which they were preferentially expressed (Figure 1n, Table S3). Veins (C_V) had the largest number of preferentially expressed genes (972), whereas bundle sheath cells (C_{BS}) had the fewest number of preferentially expressed genes (285). Functional enrichment analysis showed that genes in the mesophyll (C_M) were over-represented in photosynthesis, coenzyme metabolism and solute transport, whereas genes in C_{BS} were enriched in solute transport, enzyme classification, amino acid metabolism and polyamine metabolism (Figure 1o). Genes in C_V were involved in cellular respiration, polyamine metabolism, carbohydrate metabolism and phytohormone action (Figure 1o). $C_{BS\&M}$ contained genes highly expressed in both mesophyll and bundle sheath cells, and was over-represented in processes including photosynthesis, coenzyme metabolism, carbohydrate metabolism, redox homeostasis and secondary metabolism (Figure 1o). $C_{BS\&V}$ contained genes associated with protein biosynthesis, cellular respiration and solute transport (Figure 1o). No enriched categories were associated with both the mesophyll and the vein cells ($C_{M\&V}$). Consistent with their distinct function, vein and mesophyll clusters showed low overlap, but the most abundant transcripts in bundle sheath cells were also either expressed in veins or mesophyll cells. Overall, and associated with their morphology, the data reveal a gradient in photosynthesis-related transcripts from low in veins to high in mesophyll cells.

Patterning of gene expression in the rice bundle sheath conditions the cells for water transport

To understand the gene classes responsible for the enrichment of the transport term in the bundle sheath, we examined the expression of major transporter families in each of the six clusters. The family most enriched in C_{BS} was the major intrinsic protein (MIP) group, but multiple genes belonging to the cation diffusion facilitators superfamily, major facilitator superfamily, ion transporter superfamily, multidrug/oligosaccharidyl lipid/polysaccharide flippase superfamily and amino acid/polyamine/organocation superfamily were also present (Figure S3a).

The MIP group contains genes encoding water channels (aquaporins), including plasma membrane intrinsic proteins (PIPs) and tonoplast intrinsic proteins (TIPs) (Sakurai *et al.*, 2005). Transcripts encoding three and two members of the PIP1 and PIP2 families, respectively, accumulated preferentially in the bundle sheath compared with mesophyll and veinal cells (Figure 2a). Consistent with the deep sequencing data, *in situ* RNA localization of *PIP1.1* and *PIP1.3* generated a strong signal associated with the periphery of bundle sheath cells (Figure 2b). Transcripts encoding three tonoplast intrinsic proteins (*TIP1.1*, *TIP1.2* and *TIP2.2*) also accumulated preferentially in bundle sheath cells (Figure 2a), presumably allowing water storage in the large vacuole. Lastly, in addition to these transporters, specific P-type and V-type ATPases that could regulate leaf hydraulic conductance (Grunwald *et al.*, 2021), and establish a proton gradient across plasma and vacuole membranes to power secondary transport, were strongly expressed in bundle sheath cells (Figure S3b). Taken together, this preferential patterning of multiple aquaporins to the rice bundle sheath suggests an important role for these cells associated with water transport and storage.

© 2021 The Authors.

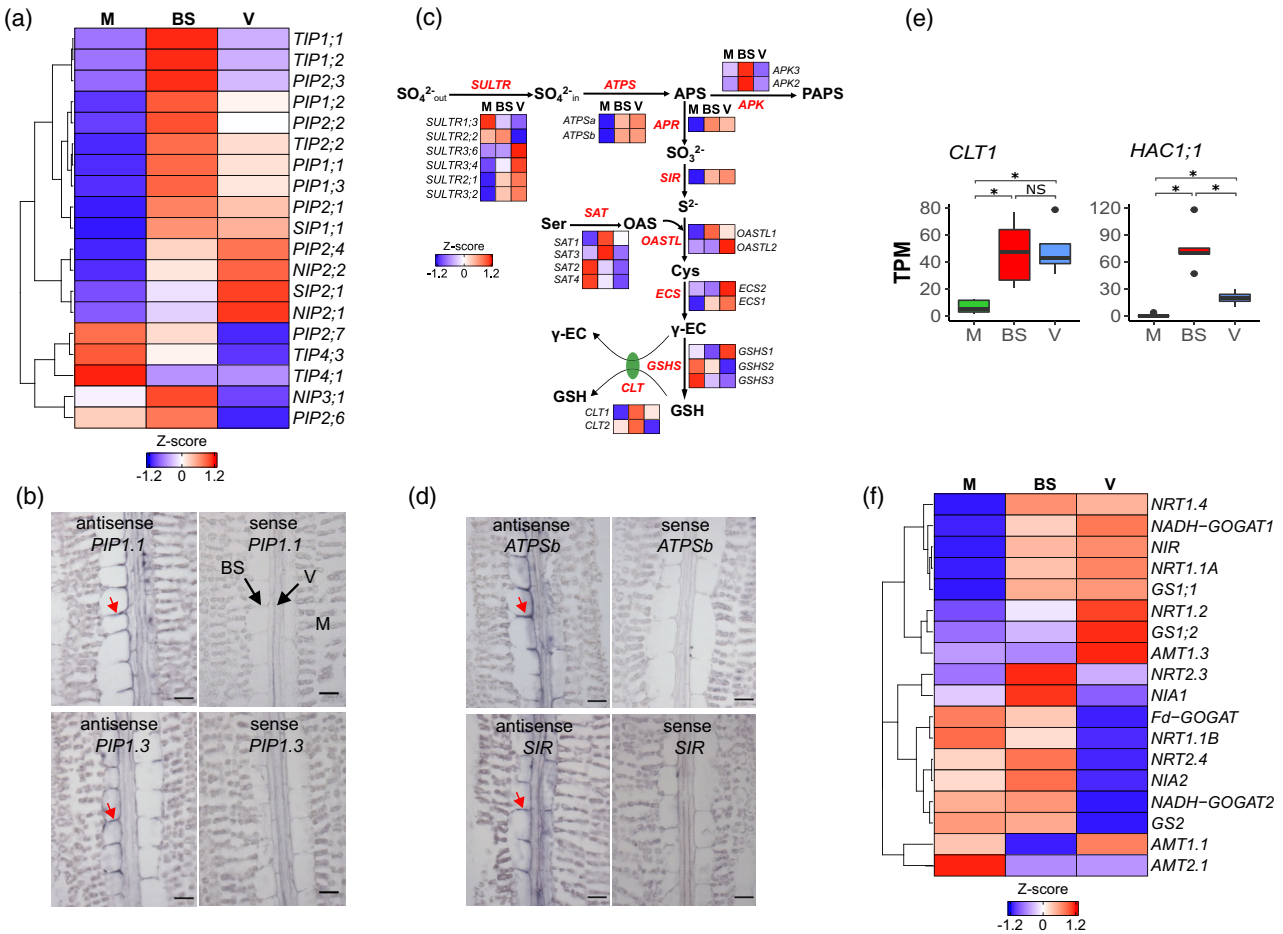


Figure 2. Preferential accumulation of transcripts associated with water transport and sulphur and nitrogen assimilation in the rice bundle sheath. (a) Relative transcript abundance for aquaporins in mesophyll, bundle sheath and veinal cells. Log_2-transformed quantile-normalized transcripts per million (TPM) values were scaled and genes with similar expression patterns were clustered using the hierarchical method. (b) Representative image after *in situ* hybridization localization for *PIP1.1* and *PIP1.3* mRNAs. Scale bars: 20 μm. Arrows indicate specific signals on bundle sheath cell periphery. (c) Schematic illustrating sulphur assimilation and relative transcript abundance in mesophyll (M), bundle sheath (BS) and veinal (V) cells, depicted as *Z*-scores from log_2-transformed quantile-normalized TPM values. (d) Representative image after *in situ* hybridization for *ATPSb* and *SIR* mRNAs. Scale bars: 20 μm. Arrows indicate specific signals on bundle sheath cell periphery. (e) Transcript abundance of *CLT1* and *HAC1;1*. (f) Relative transcript abundance for genes involved in nitrogen assimilation.

The rice bundle sheath preferentially accumulates transcripts associated with sulphur and nitrogen assimilation

Sulphur is an essential element required for both central metabolism and responding to biotic and abiotic stress. Although primarily taken up as sulphate by the roots, reduction mainly takes place in the leaves (Figure 2c). Prior to activation by ATP sulfurylase (ATPS) to adenosine 5′-phosphosulfate (APS), transport into the cell is mediated by SULTR transporters. APS is reduced into sulfite and sulfide by APS reductase (APR) and sulfite reductase (SIR), respectively, and then incorporated into *O*-acetylserine via *O*-acetylserine (thiol)lyase (OASTL) to generate the amino acid cysteine. Notably, transcripts derived from two highly expressed *SULTR* transporters (*SULTR2;1* and *SULTR3;2*), both *ATPS* genes, *APR*, as well as *SIR* and *OASTL1*, were more abundant in bundle sheath cells compared with mesophyll cells (Figure 2c). With the exception of the transcripts encoding *APR* and *OASTL1* that were most abundant in the bundle sheath cells, most were even more highly expressed in the veinal cells (Figure 2c). RNA *in situ* hybridization for transcripts encoding ATPS and SIR showed stronger signals in bundle sheath and veinal cells, compared with mesophyll cells (Figure 2d). These results strongly imply that gene expression of the rice bundle sheath cells is conditioned to allow sulphur assimilation.

The data also indicate that gene expression in the bundle sheath is conditioned for the synthesis of glutathione, a major sulphur-containing metabolite that plays critical roles in redox homeostasis and heavy metal(loid) detoxification. The biosynthesis of glutathione is catalysed by γ-glutamylcysteine synthetase (ECS) to generate γ-

© 2021 The Authors.

glutamylcysteine (γ-EC) from glutamate and cysteine, followed by the ligation of glycine and γ-EC by glutathione synthetase (Foyer and Noctor, 2011; Hernández *et al.*, 2015). The intermediate γ-EC has to be exported from the plastid by the CRT-like transporter (CLT) (Hernández *et al.*, 2015; Maughan *et al.*, 2010; Yang *et al.*, 2016) to sustain glutathione biosynthesis in the cytosol (Pasternak *et al.*, 2007). Interestingly, we found that transcripts encoding ECS1 were preferentially expressed in bundle sheath and veinal cells (Figure 2c), and that γ-glutamylcysteine transporter *CLT1* accumulated preferentially in bundle sheath cells (Figure 2c). Rice absorbs both arsenate and arsenite by different transporters, but arsenate needs to be reduced into arsenite before it can be detoxified by phytochelatin. *CLT1* has been reported to be critical for rice tolerance to arsenic because it determines phytochelatin biosynthesis and arsenate reduction (Yang *et al.*, 2016). Notably, the arsenate reductase *HAC1;1* also showed preferential expression in the bundle sheath, suggesting that this cell type may play an important role in arsenate reduction and detoxification (Figure 2e).

As with sulphur assimilation, transcripts encoding some of the pathway allowing nitrate reduction were more highly expressed in the bundle sheath and veinal cells compared with the mesophyll cells. Interestingly, this included the nitrate transporters *NRT1.4, NRT1.1A*, *NRT1.2* and *NRT2.3*, both nitrate reductases (*NIA1* and *NIA2*), nitrite reductase (*NIR*) and glutamine synthetase (*GS1.1*). Transcripts encoding *NRT2.3*, *NIA1* and *NIA2* were most abundant in bundle sheath cells, whereas the rest were also highly expressed in veins. In contrast, transcripts encoding glutamine synthetase (*GS2*) and glutamate synthase (*Fd-GOGAT*) that allow ammonia assimilation in the chloroplast were preferentially expressed in the bundle sheath and the mesophyll cells, relative to the veinal cells (Figures 2f and S4). These results indicate that gene expression in the rice bundle sheath is also tuned to specialize in nitrate assimilation and amino acid biosynthesis.

The Calvin–Benson–Bassham cycle, photorespiration and C_4 cycle gene expression

Compared with the mesophyll cells, bundle sheath cells contain few chloroplasts and veinal cells only contain rudimentary plastids (Figure 3a). Consistent with this, transcripts encoding most components of the photosynthetic electron transport chain were abundant in mesophyll cells, and although barely detectable in veins they were clearly expressed in the bundle sheath (Figure S5a). Exceptions included one ferredoxin gene that showed the highest transcript abundance in bundle sheath cells, three ferredoxin genes for which transcripts were more abundant in bundle sheath and veinal cells, compared with mesophyll cells, and three homologues of photosystem I subunit PsaA and photosystem II subunit PsbK (Figure S5a, b). In contrast to the primary ferredoxin Fd1 (LOC_Os08g01380) involved in photosynthetic electron transport (He *et al.*, 2020), and primarily expressed in mesophyll cells, four other ferredoxins including the nitrate-inducible Fd (LOC_Os05g37140) (Doyama *et al.*, 1998) were preferentially expressed in bundle sheath and veinal cells, indicating that rice bundle sheath cells have optimised reducing power for nitrate reduction as well as sulphur assimilation. A gradient from high expression in mesophyll cells to low expression in veinal cells was observed for most enzymes of the Calvin–Benson–Bassham cycle. Exceptions included rubisco activase (*RCA*), which was poorly expressed in both bundle sheath and veinal cells, and RBCS4, which had similar transcript abundance in bundle sheath and mesophyll cells (Figures 3b and S5c). The ratio of *RCA* to *RbcS* transcripts was twofold higher in mesophyll cells compared with bundle sheath cells. With the exception of ADP-glucose pyrophosphorylase subunits (APL1 and APS2), starch synthase (SSI) and granule-bound starch synthase (GBSSII), which were more abundant in mesophyll cells, transcripts encoding enzymes of starch synthesis were similar in mesophyll and bundle sheath cells (Figure S5d). Although transcripts associated with the glucose 6-phosphate/phosphate translocator (GPT1 and GPT2-1) were more abundant in bundle sheath cells than in mesophyll cells, their abundance was low compared with the triose phosphate/phosphate translocator (TPT1) (Figure S5d). Together, these data indicate that gene expression in the bundle sheath is set up to favour starch synthesis using products of the Calvin–Benson–Bassham cycle.

Most transcripts encoding proteins of photorespiration showed strong expression in the mesophyll, weaker expression in the bundle sheath, and very poor expression in veins (Figures 3c and S5e). The only exceptions were genes with very low absolute levels of expression that were preferential to veins, and which included one L-glycine decarboxylase subunit (*GDCL2*), one serine hydroxymethyltransferases (*SHMT4*), one phosphoglycolate phosphatase (*PGP2*) and two glycine decarboxylase complex H subunits (*GDCH2* and *GDCH3*) (Figures 3c and S5e).

It has previously been reported that vascular tissue of C_3 plants possesses high activities of some enzymes associated with C_4 photosynthesis, including all three C_4 acid decarboxylases and pyruvate, orthophosphate dikinase (Brown *et al.*, 2010; Hibberd and Quick, 2002; Shen *et al.*, 2016). However, to our knowledge it has not been possible to delimit these activities to the specific cells associated with the vascular tissue. To investigate this, we assessed the abundance of transcripts encoding core enzymes of the C_4 cycle in the veinal, bundle sheath and mesophyll cells of rice (Figure 3d). Transcripts of carbonic anhydrase (*CA*) were more abundant in the mesophyll cells compared with the bundle sheath and veinal cells, a pattern consistent with that required for C_4 photosynthesis. Transcripts

© 2021 The Authors.
The Plant Journal published by Society for Experimental Biology and John Wiley & Sons Ltd,

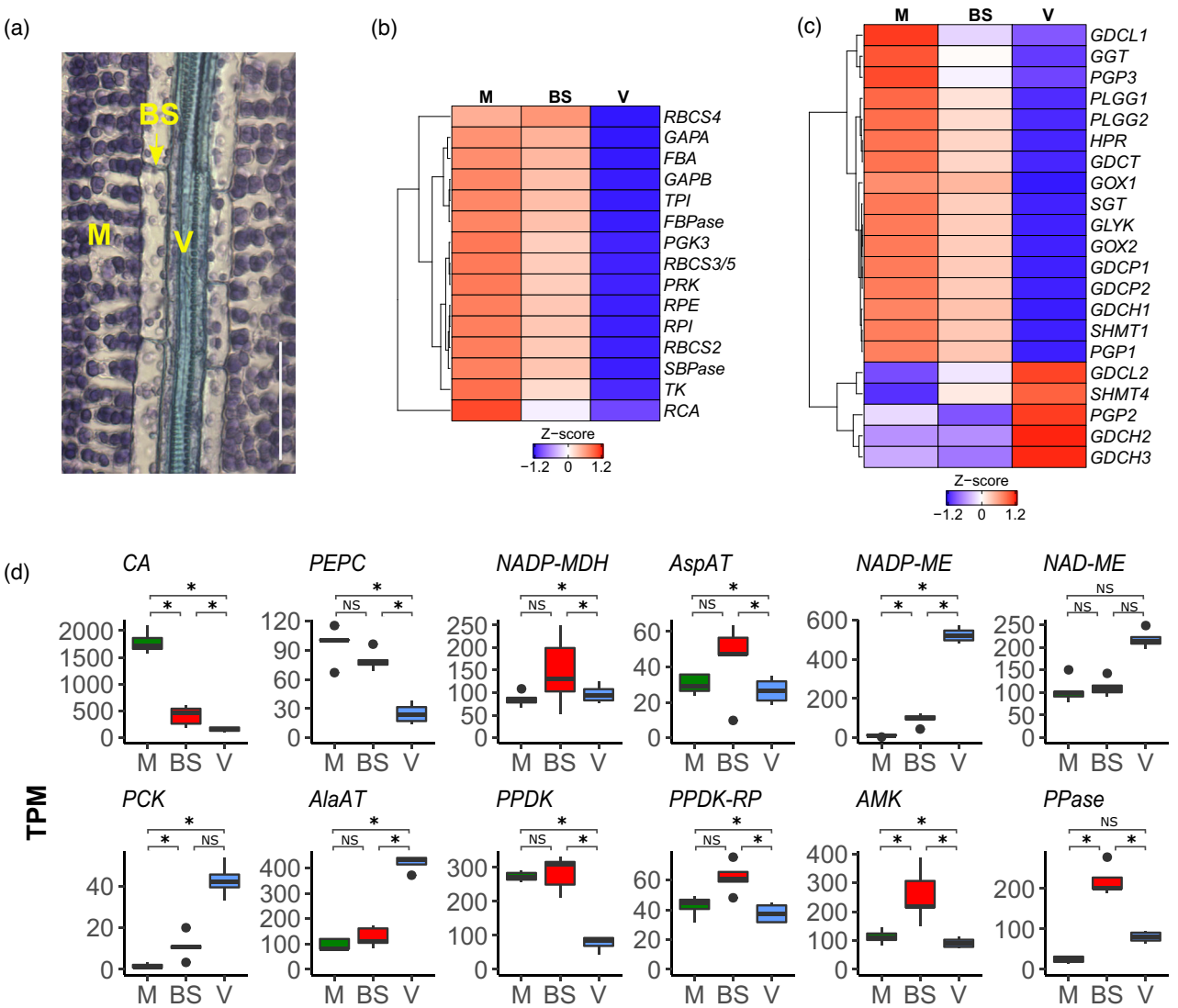


Figure 3. (a) Paradermal section of rice leaf stained with toluidine blue shows that bundle sheath cells are less occupied by chloroplasts compared with mesophyll cells. Scale bar: 50 μm. Relative transcript abundance for genes involved in the Calvin–Benson–Bassham cycle (b), photorespiration (c) and C_4 pathway (d), in mesophyll (M), bundle sheath (BS) and veinal (V) cells of *Oryza sativa* (rice). (b, c) Transcripts associated with Calvin–Benson–Bassham cycle (b) and photorespiration (c) were preferentially expressed in mesophyll cells. Log_2-transformed quantile-normalized transcripts per million (TPM) values were scaled and genes with similar expression pattern were clustered using the hierarchical method. (d) Transcripts encoding the C_4 acid decarboxylases NADP-dependent malic enzyme (NADP-ME) and phosphoenolpyruvate carboxykinase (PCK) accumulated preferentially in bundle sheath and veinal cells, whereas ancillary enzymes AMP kinase (AMK) and pyrophosphorylase (PPASE) for pyruvate orthophosphate dikinase (PPDK) accumulated preferentially in bundle sheath cells. Data are presented as TPM values; asterisks indicate statistically significant difference (FDR and adjust $P < 0.05$ using EDGER and DESEQ2 analysis).

encoding PEP carboxylase (*PEPC*), NADP-malic dehydrogenase (*NADP-MDH*), aspartate aminotransferase (*AspAT*), NAD-malic enzyme (*NAD-ME*), alanine aminotransferase (*AlaAT*), pyruvate, orthophosphate dikinase (*PPDK*) and the PPDK regulatory protein (*PPDK-RP*) showed no significant difference in abundance between mesophyll and bundle sheath cells. However, transcripts encoding AMP KINASE (*AMK*) and pyrophosphorylase (*PPASE*), which allow the PPDK reaction and APS synthesis by ATPS to proceed, showed higher abundance in the bundle sheath cells than in both veinal and mesophyll cells (Figure 3d), suggesting that the activity of PPDK might be higher in this cell type. Notably, two C_4 acid decarboxylases, NADP-dependent malic enzyme (*NADP-ME*) and phosphoenolpyruvate carboxykinase (*PCK*), showed a gradient in transcript abundance from veinal to bundle sheath to mesophyll cells. These data indicate that the high activity of these C_4 decarboxylases in the vascular bundles of rice (Shen *et al.*, 2016) is likely to be caused by their expression in veinal rather than bundle sheath cells. Overall, C_4 genes could be partitioned into three main groups, either showing a strong negative gradient (*CA* and *PEPC*) from mesophyll to vein

© 2021 The Authors.

cells, a strong positive gradient (*NADP-ME*, *NAD-ME*, *PCK* and *AlaAT*) or a tendency to be most strongly expressed in the bundle sheath (*NADP-MDH*, *AspAT*, *PPDK*, *PPDK-RP*, *AMP* and *PPase*) (Figure 3d). This finding is consistent with them being part of multiple gene regulatory networks in the ancestral C_3 state that pattern their expression across these cell types.

Transcription factors and their cognate *cis*-elements associated with the rice mesophyll, bundle sheath and vein

To gain insight into the regulatory architecture associated with the three cell types, we identified transcription factors in each of the six clusters and designated these as TF_M, TF_{BS}, TF_V, $TF_{BS\&M}$, $TF_{BS\&V}$ and $TF_{M\&V}$. From a total of 201 differentially expressed transcription factors, over 30% of them (66) showed vein-specific expression (TF_V, Figure 4a), including families such as bZIP, bHLH, G2-like, MYB-related and Dof (Figure S6a). ERF, HD-ZIP and MYB-related transcription factors were the most abundant TF families in the mesophyll-specific cluster TF_M (Figures 4a and S6a). Transcription factors known to regulate chloroplast biogenesis and photosynthesis were most abundant in mesophyll cells, but were also detected in the bundle sheath. This included GNC (LOC_Os06g37450) in TF_M, and GLK1 (LOC_Os06g24070), GLK2 (LOC_Os01g13740) and CGA1 (LOC_Os02g12790) in $TF_{BS\&M}$. Thus, consistent with the chloroplast complement of each cell type (Figure 3a), these transcription factors found in $TF_{BS\&M}$ were less abundant in the bundle sheath compared with mesophyll cells (Figure S6c).

The bundle sheath-specific cluster contained only 10 genes that derived from families such as the ERFs, bZIPs and bHLHs (Figures 4a and S6a). Transcription factors abundant in both bundle sheath and vein belonged to the MYB, ERF, bZIP and WRKY families, whereas CO-like, C_2H_2 and G2-like families were abundant in the mesophyll and bundle sheath cluster (Figures 4a and S6a). The ZF-HD, G2-like, DBB, MIKC-MADS and Dof families were significantly enriched in the vein-specific cluster, the CO-like transcription factors were over-represented in $TF_{BS\&M}$, and the Dof, MYB, HRT-like and VOZ transcription factors were over-represented in $TF_{BS\&V}$ (Figure S6b).

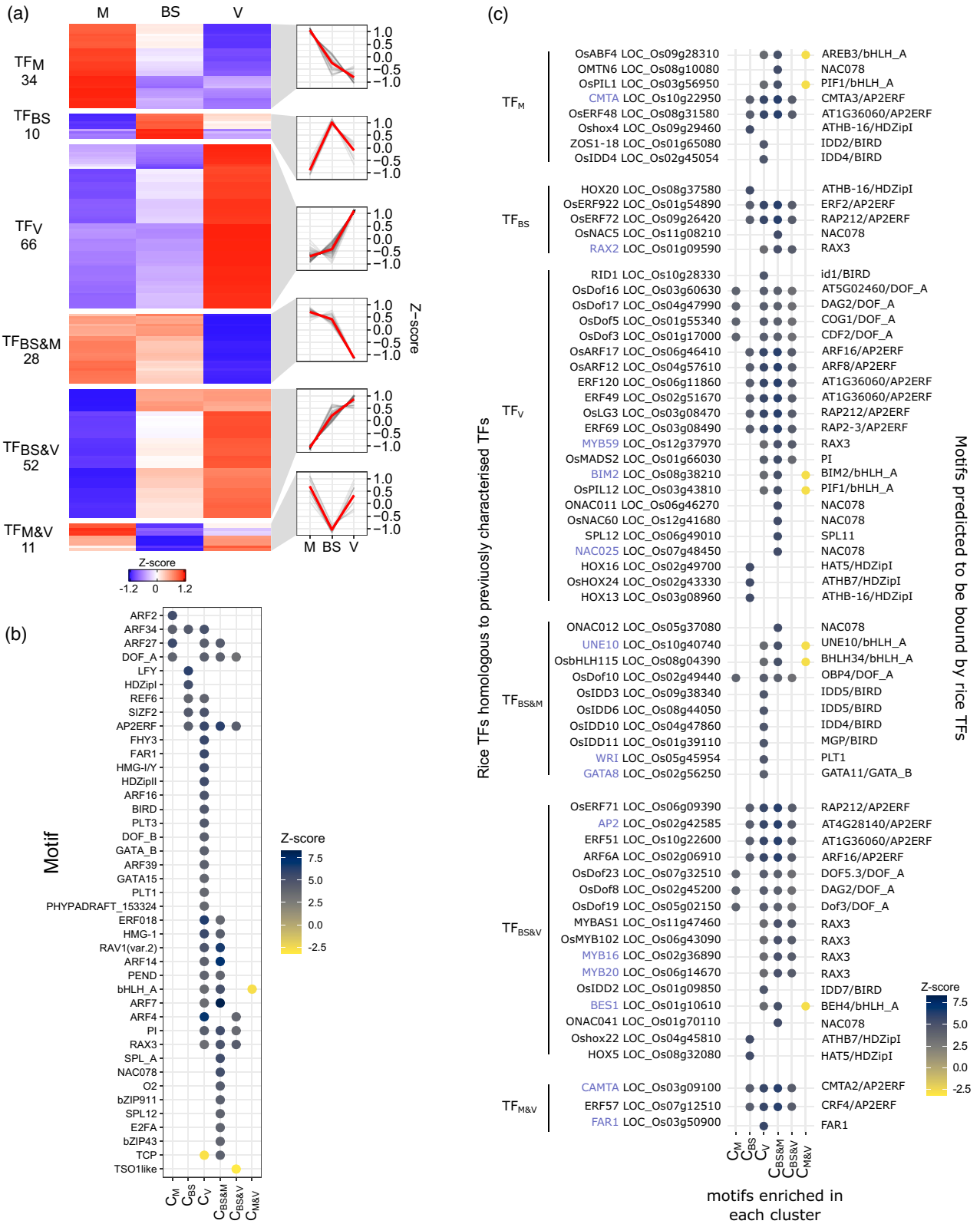
To determine whether DNA motifs known to be bound by transcription factors were associated with genes preferential to each cell type, we performed motif enrichment analyses using the FIMO and AME tools (Bailey *et al.*, 2015). Although the two approaches differ in their statistical testing they returned broadly similar estimates for the number of enriched motifs. AME found the highest number of motifs in C_{BS} and $C_{BS\&M}$ (n = 44 and 91, respectively), whereas FIMO identified the most enriched motifs in C_V and $C_{BS\&M}$ (n = 86 and 100, respectively) (Tables S4 and S5). Both found either few or no enriched motifs in the $C_{BS\&V}$ (AME = 11; FIMO = 5), C_M (AME = 0; FIMO = 3) or $C_{M\&V}$ (both 0) clusters (Table S4), suggesting that genes making up these three clusters are heterogeneously regulated.

As many DNA motifs share considerable sequence similarity, and closely related transcription factors can bind to similar motifs, we collapsed known DNA binding motifs from closely related transcription factors into 41 groups (Figure 4b). Twenty-three of these groups were enriched or depleted in a specific cluster associated with transcript abundance in the three cell types. Although many were found in multiple clusters, there were no examples where a particular motif was statistically enriched or depleted in all six clusters (Figure 4b). C_M contained the fewest enriched motifs (n = 4), only one of which (ARF2) was uniquely enriched in this cluster (Figure 4b). HD-Zip I and LHY motifs were unique to C_{BS}, but all other enriched motifs in the bundle sheath were also found in C_V, suggesting an overlap in the regulation of gene expression between bundle sheath and veinal cells (Figure 4b). C_V contained the most enriched motifs (n = 29), of which almost half of these were unique to this cluster, and included BIRD, Dof, GATA, HD-Zip II, PLT, FAR1 and FHY3 motifs (Figure 4b).

We next investigated whether transcription factors and their cognate DNA binding sites were associated with the six clusters that had been defined by patterns of gene expression across mesophyll, bundle sheath and veinal cells. As relatively few transcription factors from rice have

Figure 4. Patterning of transcription factors (TFs) between mesophyll, bundle sheath and veinal cells of *Oryza sativa* (rice). (a) TFs from clusters C_M, C_{BS}, C_V, $C_{BS\&M}$, $C_{BS\&V}$ and $C_{M\&V}$ were designated as TF_M, TF_{BS}, TF_V, $TF_{BS\&M}$, $TF_{BS\&V}$ and $TF_{M\&V}$, respectively, and the relative abundances of differentially expressed TFs were presented as a heatmap. A line plot of the *Z*-score, calculated from $\log_2$-transformed quantile-normalized transcripts per million (TPM) values, is presented; the red lines in these plots represent the mean values of the *Z*-score. (b) Significantly enriched or depleted motifs were identified in each of the cistromes from the six gene expression clusters. Enrichment was calculated using the REGIONER permutation testing package (Gel *et al.*, 2016) following motif scanning using FIMO to identify motifs from the plant Jaspar non-redundant database (Fornes *et al.*, 2020). The *Z*-scores are shown with a colour scale to show the magnitude of enrichment (dark blue) or depletion (yellow) for motifs that were significant after multiple testing correction. Motifs derived from closely related TFs were grouped together for visualization based on their degree of overlap with predicted target sites (e.g. AP2ERFs). The cistrome from cluster C_V shows the greatest number of enriched motifs, including 13 uniquely enriched motifs, whereas the C_M and C_{BS} cistromes have far fewer enriched motifs. (c) Cluster-specific TFs (left of panel) were mapped to the motifs (right of panel) that they are most likely to bind to, based on high protein sequence similarity with the proteins in the Jaspar plant motif database. The TFs that mapped to any enriched motifs are shown with the motif enrichment data. This allows the visualization of the intersection between TF transcript abundance and potential activation activity. Gene abbreviations for rice TFs were retrieved from the funRiceGene database (Yao *et al.*, 2018), but for abbreviations not found in the database the symbols of best-hit Arabidopsis TFs were used and presented in blue. The matching motifs show first the best match and then the motif group if part of a group as shown in (b).

© 2021 The Authors.
The Plant Journal published by Society for Experimental Biology and John Wiley & Sons Ltd,



© 2021 The Authors.
The Plant Journal published by Society for Experimental Biology and John Wiley & Sons Ltd,

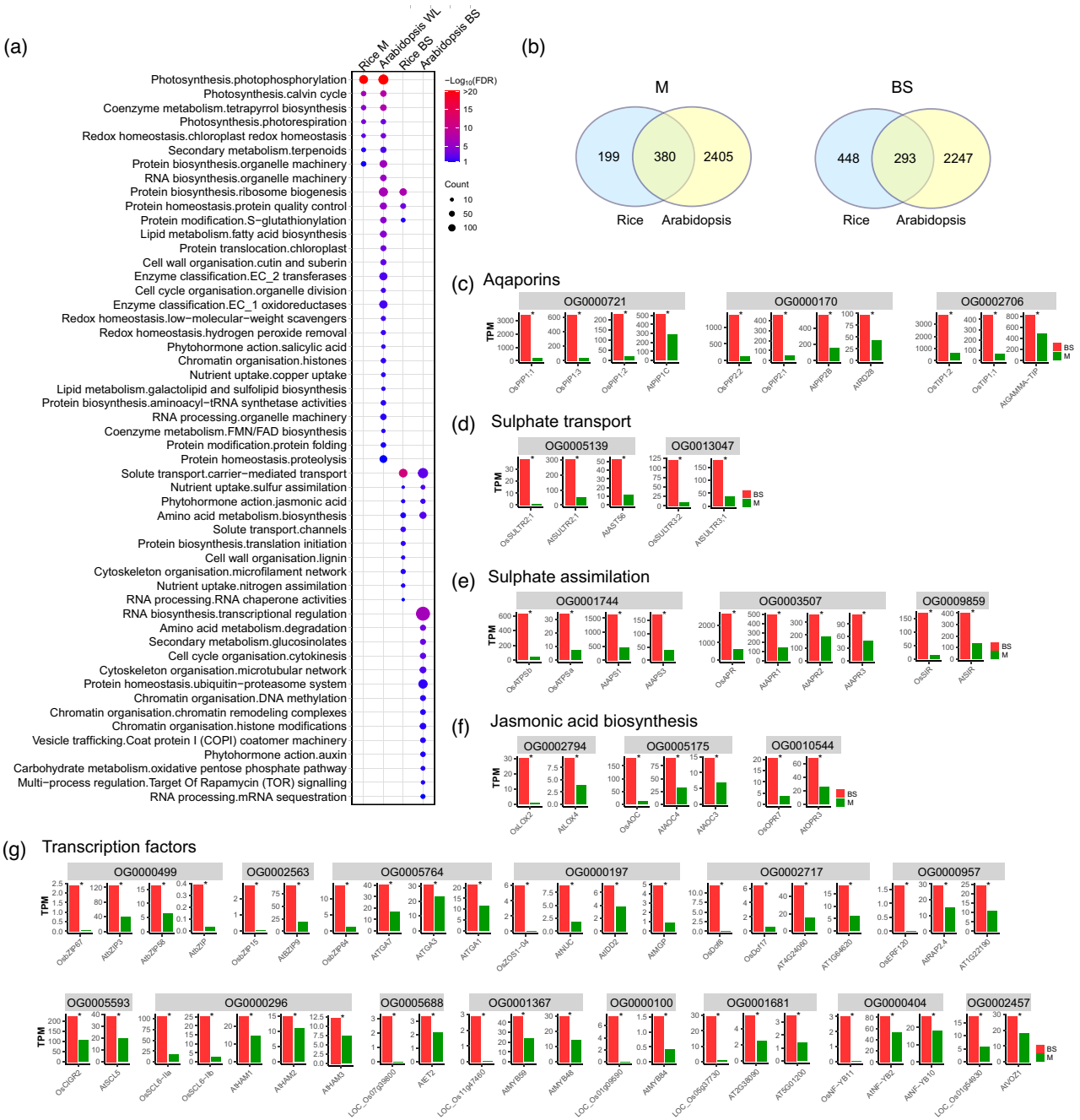


Figure 5. Conserved patterning of gene expression in the Arabidopsis and *Oryza sativa* (rice) bundle sheath. Orthologues from Arabidopsis and rice associated with aquaporins, sulphate transport and assimilation, as well as jasmonic acid biosynthesis, are strongly expressed in bundle sheath cells. (a) The enriched MAPMAN categories (secondary level) of bundle sheath (BS) and mesophyll (M) preferential genes in rice and Arabidopsis were defined using Fisher's exact test (FDR < 0.1). (b) Venn diagram illustrating the extent to which genes in the same orthogroup are preferentially expressed in mesophyll or bundle sheath cells of both rice and Arabidopsis. (c–g) Transcript abundance of Arabidopsis and rice genes belonging to the same orthogroups of aqaporins (c), sulphate transport (d), sulphate assimilation (e), jasmonic acid biosynthesis (f) and transcription factors (g). Data are presented as transcripts per million (TPM) and statistically significant differences are annotated with an asterisk (FDR and adjusted $P < 0.05$ using EDGER and DESEQ2 analysis in this study, posterior probability of differential expression > 0.95; Aubry *et al.*, 2014b); red and green bars represent bundle sheath and mesophyll, respectively.

had their DNA binding characteristics determined, we first used protein homology (BLASTP bit-score > 100) to link them with transcription factors for which DNA binding data are available. Of the mesophyll-specific transcription factors, although five are likely to bind motifs enriched in the cistrome of bundle sheath and mesophyll cells (Figure 4c), none were associated with motifs enriched only in the mesophyll cistrome. In contrast, three bundle sheath-specific

© 2021 The Authors.
The Plant Journal published by Society for Experimental Biology and John Wiley & Sons Ltd,

transcription factors coincided with motifs enriched in the BS-specific cistrome. Vein-specific transcription factors showed the greatest convergence with enrichment in their cognate DNA binding sites, with 15 mapping to the vein-specific cistrome. This included transcription factors predicted to bind BIRD, Dof, AP2ERF, MYB, MADS and bHLH motifs (Figure 4c). Overall, this approach identifies families of transcription factors preferentially expressed in cell types in which their cognate motifs were over-represented in the cistrome of that cell type. We regard these transcription factors as strong candidates for patterning gene expression across these cell types.

Conserved patterning of gene expression in bundle sheath cells from rice and Arabidopsis

Bundle sheath cells are present in both monocotyledons and dicotyledons. Previous analysis has compared transcripts loaded onto ribosomes in the bundle sheath of Arabidopsis with those from whole leaves. In so doing, it was concluded that the bundle sheath in Arabidopsis is important for sulphur and glucosinolate metabolism as well as trehalose synthesis (Aubry *et al.*, 2014b). The functional enrichment of bundle sheath and mesophyll preferential genes from rice and Arabidopsis were compared. Consistent with the mesophyll representing the major cell type conducting photosynthesis, gene sets common to the mesophyll from the two species included transcripts important for the photosynthetic electron transport chain and photophosphorylation, the Calvin–Benson–Bassham cycle, photorespiration, tetrapyrrole biosynthesis, chloroplast redox homeostasis, organelle protein biosynthesis and terpenoid biosynthesis (Figure 5a). In contrast, only four categories of genes were enriched in the bundle sheath from both species, namely carrier-mediated transport, sulphur assimilation, amino acid biosynthesis and jasmonic acid action (Figure 5a).

We next identified genes from both species likely to be descended from a common ancestor and placed these into orthogroups (Emms and Kelly, 2019). Genes were assigned to 10 665 orthogroups and the extent to which each was preferentially expressed in mesophyll or bundle sheath cells determined. Genes from 380 orthogroups were preferentially expressed in mesophyll cells of rice and whole leaves of Arabidopsis, whereas genes from 293 orthogroups were shared by bundle sheath of both species (Figure 5b), in both cases with a greater overlap than would be expected by chance (Fisher's exact test, $P < 2.2e^{-16}$ for both bundle sheath and mesophyll). However, the odds ratio that is also generated from the Fisher's exact test indicated that there was a greater degree of overlap between the mesophyll of rice and whole leaves of Arabidopsis than between their bundle sheath cells (odds ratios were 5.32 and 2.06 for mesophyll and bundle sheath, respectively). We thus conclude that the expression of orthologs has diverged more in the bundle sheath than in the mesophyll of these species. Specific orthogroups associated with the mesophyll in both species were mainly associated with photosynthesis but also included 11 orthogroups of transcription factors, one of which included the transcription factor GLK1 (OG0002393) (Table S5) that is known to activate photosynthesis-related gene expression (Waters *et al.*, 2008; 2009). Although fewer orthogroups were common to the bundle sheath in rice and Arabidopsis, overlap included three groups encoding aquaporins (PIP1, OG0000721; PIP2, OG0000170; and TIP1, OG0002706), as well as orthologs involved in sulphate transport (OG0005139 and OG0013047), sulphate assimilation (OG0001744, OG0003507 and OG0009859) and jasmonic acid biosynthesis (OG0002794, OG0005175 and OG0010544) (Figures 5c–f and S7, Table S5). We compared bundle sheath preferentially expressed genes of rice and Arabidopsis with the bundle sheath marker genes identified through single-cell sequencing of Arabidopsis (Kim *et al.*, 2021). As the sequencing depth from single-celled sequencing is not as great, fewer transcripts were detected (Figure S8a). However, 41 orthogroups were identified containing at least one gene with preferential expression in the bundle sheath in all three studies, including two sulphate transporters (OG0005139 and OG0013047), ATPS (OG0001744) and allene oxide cyclase (OG0005175). To investigate whether these genes are associated with bundle sheath cells in a broader range of species, we assessed transcript abundance of consensus orthogroups associated with water transport, sulphur assimilation, jasmonic acid biosynthesis as well as nitrate reduction in publicly available data from *Panicum virgatum* (Rao *et al.*, 2016), *Setaria italica* (John *et al.*, 2014), *Sorghum bicolor* (Emms *et al.*, 2016) and *Z. mays* (Chang *et al.*, 2012), as well as the C_4 dicotyledon *Gynandropsis gynandra* (Aubry *et al.*, 2014a). Except for nitrate reduction genes, which are more highly expressed in mesophyll cells, this indicated that in most of these species the genes in each of the orthogroups are preferentially expressed in bundle sheath cells (Figure S9). Taken together, the data strongly suggest that sulphur assimilation, jasmonic acid biosynthesis and water transport represent ancestral functions associated with the bundle sheath derived from the last common ancestor of monocotyledons and dicotyledons.

We next wished to identify whether any transcription factors were preferentially expressed in the bundle sheath cells of both species. The rice orthologue of the key regulator of the sulphur starvation response sulfur limitation 1 (Maruyama-Nakashita *et al.*, 2006), *OsEIL3* (LOC_Os09g31400), showed low expression but, consistent with Arabidopsis, was more strongly expressed in bundle sheath compared with mesophyll cells (Figure S7). Fourteen transcription factor orthogroups were identified containing at least one orthologue in both species that was

© 2021 The Authors.

preferential to the bundle sheath (Figure 5g). These included three basic leucine zipper (OG0000499, OG0002563 and OG0005764), C_2H_2-type zinc finger (OG0000197), DNA-binding with one finger (OG0002717), ethylene-responsive factor (OG0000957), two GRASs (OG0000296 and OG0005593), hairy-related transcription factor (OG0005688), three MYBs (OG0000100, OG0001367 and OG0001681), nuclear factor-YB (OG0000404) and vascular plant one zinc finger protein (OG0002457) transcription factor families (Figures 5g and S7; Table S5). These data imply purifying selection has acted to maintain the expression of these transcription factors in the bundle sheath, as these species diverged from their last common ancestor prior to the divergence of the dicotyledons and monocotyledons.

DISCUSSION

Patterning of photosynthesis gene expression between cell types in the rice leaf

The separation of protoplasts followed by cell sorting has provided significant insight into gene expression in specific cells of roots (Birnbaum *et al.*, 2003; Brady *et al.*, 2007). In leaves, the ability to separate mesophyll and bundle sheath cells from C_4 plants (Moore *et al.*, 1984; Edwards and Black, 1971; Kanai and Edwards, 1973) has allowed similar levels of insight into these cells, but leaf cells from C_3 species have been more challenging to isolate. As a consequence, although the major cell types of leaves such as the mesophyll, phloem, xylem and guard cells have well-defined roles, others such as the bundle sheath are relatively poorly understood (Leegood, 2008). In Arabidopsis the bundle sheath is thought to play important roles in regulating hydraulic conductance, substrate transport and storage (Griffiths *et al.*, 2013; Sade *et al.*, 2014; Shatil-Cohen *et al.*, 2011). Analysis of mRNAs resident on ribosomes showed that patterns of gene expression in the Arabidopsis bundle sheath are conditioned to facilitate sulphur metabolism and glucosinolate biosynthesis (Aubry *et al.*, 2014b). Although the suppression of chlorophyll synthase in veinal tissue including bundle sheath cells of Arabidopsis reduced photosynthesis, growth and fitness (Janacek *et al.*, 2009), the importance of the bundle sheath itself was not defined. Moreover, in C_3 monocotyledons, the group containing many of our most important crops, little is known about the function of bundle sheath cells. To address this, we defined gene expression in the rice bundle sheath as well as mesophyll and veinal cells using LCM coupled with mRNA sequencing. Contrary to our expectations, this indicated that although the bundle sheath cells of rice contain few chloroplasts (Sage and Sage, 2009; Wang *et al.*, 2017), transcripts encoding components of the photosynthetic electron transport chain and the Calvin–Benson–Bassham cycle were clearly expressed in rice bundle sheath cells. The consequence of this finding is that the expression of photosynthesis genes represents a continuum from low to medium to high in veinal, bundle sheath and mesophyll cells, respectively. Whether this is also the case in Arabidopsis remains to be determined. There is significant interest in activating photosynthesis in the bundle sheath of rice, such that mesophyll and bundle sheath cells could be engineered to carry out C_4 photosynthesis (Sage, 2004; Wang *et al.*, 2017). The finding that photosynthesis gene expression in the bundle sheath resembles the mesophyll more than veins indicates that it needs to be retuned, rather than completely reprogrammed, to achieve this demanding aim.

Expression of C_4 genes in the C_3 rice leaf

Vascular bundles of C_3 plants are known to carry out C_4-like metabolism via the high activities of C_4 acid decarboxylases and PPDK that make use of the C_4 acids present in the transpiration stream (Brown *et al.*, 2010; Hibberd and Quick, 2002; Shen *et al.*, 2016). To date, it has not been possible to define whether the high activity of these C_4 enzymes in C_3 plants is associated with their expression in veinal or bundle sheath cells. The analysis of rice we present here shows that transcripts encoding two C_4 acid decarboxylases, NADP-ME and PCK, were more abundant in bundle sheath cells compared with mesophyll cells, but in both cases transcripts were even more strongly expressed in veinal cells. These findings suggest that the high activity of NADP-ME and PCK in vascular tissue is primarily caused by expression in the veins rather than in the bundle sheath. Thus, the gradient in expression of genes encoding these C_4 acid decarboxylases, from high in veinal cells to low in mesophyll cells, is the opposite of that for photosynthesis genes in these cell types. Whether this is also the case in C_3 dicotyledons such as Arabidopsis remains to be determined.

Although there was no significant difference in *PPDK* transcript abundance between bundle sheath cells and mesophyll cells, we detected greater abundance of transcripts encoding AMK and PPase in bundle sheath cells, which carry out ancillary reactions, allowing the PPDK reaction to proceed. As the activity of PPDK is higher in vascular strands of rice than in mesophyll cells (Shen *et al.*, 2016), it is therefore possible that this is caused by greater activity of AMP and PPase in bundle sheath cells. PPDK is important for nitrogen recycling in Arabidopsis and *Nicotiana tabacum* (tobacco) (Taylor *et al.*, 2010), and so these data also suggest that this is likely to be the case in rice. Although *CA* and *PEPC* transcripts in rice were most abundant in mesophyll cells, the specific isoforms that these transcripts encode are predicted to generate proteins localized to the chloroplast (Chen *et al.*, 2017; Masumoto *et al.*, 2010). In C_4 leaves both proteins are cytosolic (Hatch, 1987; Hatch and Burnell, 1990). Overall, these

© 2021 The Authors.

results suggest that at least three modifications to the expression of these genes would be required to build a C_4 cycle in rice: amplifying the existing expression of C_4 acid decarboxylases in bundle sheath cells; repositioning CA and PEPC proteins such that they reside in the cytosol; and expressing AMK and PPase in the mesophyll cells for effective PPDK activity.

The rice bundle sheath is conditioned to allow water transport and sulphate and nitrate metabolism

Our analysis of the rice bundle sheath indicates that gene expression is poised to allow photosynthesis in this cell type. This finding is consistent with the fact that although during early leaf development plastids of the rice bundle sheath contain significant levels of starch, as the leaf matures these plastids develop into chloroplasts (Miyake, 2016). Transcripts patterns in the bundle sheath were consistent with starch being synthesized from carbon skeletons derived from the Calvin–Benson–Bassham cycle, with strong expression of *PGI*, *PGM*, *UGP1*, *SSIIB* and *SSIII-1*, and with low expression of hexose phosphate transporters (*GPT1* and *GPT2-1*), compared with the *TPT1*. Our analysis also indicates that the rice bundle sheath is patterned to facilitate water transport and storage and sulphate and nitrate assimilation. It was notable that most of highly expressed aquaporins were preferentially expressed in bundle sheath cells, and included members of the PIP1, PIP2, TIP1 and TIP2 subfamilies. It has been reported that water transport activity and plasma membrane localization of PIP1 requires interaction with PIP2 (Fetter *et al.*, 2004; Zelazny *et al.*, 2007), and so their co-expression in the bundle sheath is compatible with efficient water transport and storage in this cell type. Notably, analysis of publicly available data indicates that strong expression of *PIP1* and *PIP2* in the bundle sheath is also found in the C_4 grasses *P. virgatum*, *Setaria italica*, *Sorghum bicolor* and *Z. mays*, as well as the C_4 dicotyledon *G. gynandra* (Aubry *et al.*, 2014a; Chang *et al.*, 2012; Emms *et al.*, 2016; John *et al.*, 2014; Rao *et al.*, 2016), implying that the bundle sheath plays an important role in water transport in the common ancestor of monocotyledons and dicotyledons (Figure S9).

Enzymes of sulphur assimilation preferentially accumulate in the bundle sheath of C_4 grasses and C_3 dicotyledon Arabidopsis (Aubry *et al.*, 2014b; Burgener *et al.*, 1998; Burnell, 1984; Chang *et al.*, 2012; Emms *et al.*, 2016; Gerwick *et al.*, 1980; John *et al.*, 2014; Passera and Ghisi, 1982; Rao *et al.*, 2016; Schmutz and Brunold, 1984), but this is not the case in wheat and the C_4 dicotyledons *Flaveria trinervia* and *G. gynandra*, where ATPS and APR showed similar activity and/or transcript abundance in bundle sheath and mesophyll cells (Aubry *et al.*, 2014a; Kopriva *et al.*, 2001; Schmutz and Brunold, 1984). We found that transcripts associated with sulphur assimilation were restricted or preferentially localized to the bundle sheath cells of rice, which is consistent with expression patterns in Arabidopsis (Aubry *et al.*, 2014b). This was also true for genes indirectly associated with sulphur assimilation, such as the PPase that increases the rate of the ATP sulphurylase reaction. Several evolutionary drivers for the localization of sulphur assimilation in the bundle sheath of C_4 grasses have been discussed, including colocalization with photorespiration, as a source of serine for cysteine synthesis, and protection of the reaction intermediates from oxidation by photosystem II-derived oxygen (Kopriva and Koprivova, 2005).

In Arabidopsis glucosinolate synthesis is controlled by a MYC–MYB transcription factor module that has recently been shown to pattern gene expression to the bundle sheath (Dickinson *et al.*, 2020). It seemed likely that the localization of glucosinolates to the bundle sheath of Arabidopsis led to the upregulation of ATPS and APR in these cells. However, the preferential expression of genes associated with sulphur assimilation in the rice bundle sheath that does not synthesize glucosinolates indicates a much more ancient role for the bundle sheath in sulphur assimilation. It also removes the proposed connection between C_4 photosynthesis and the localization of sulphur assimilation in the bundle sheath. More broadly, it appears that the bundle sheath is conditioned for sulphur assimilation in all species analysed to date. In plants such as *Triticum aestivum* (wheat), *Flaveria* and *Gynandropsis*, mesophyll cells are also used (Aubry *et al.*, 2014a; Koprivova *et al.*, 2001; Schmutz and Brunold, 1984), but in other plants the pathway appears to be restricted to the bundle sheath (Figure S9). It is not clear whether the ancestral state was for gene expression allowing sulphur assimilation to be associated with bundle sheath and mesophyll cells or whether in some species the expression in mesophyll cells has been gained. Moreover, a full complement of transporters and enzymes for nitrate reduction were found to be expressed preferentially in rice bundle sheath cells, and other enzymes that serve to provide a carbon skeleton for amino acid metabolism, such as PEPC (Masumoto *et al.*, 2010) and PPDK (Taylor *et al.*, 2010), also showed high expression in these cells. The compartmentation of nitrate and ammonia assimilation gene expression between bundle sheath and mesophyll cells of rice contrasts with their spatial separation of C_4 grasses and *G. gynandra*, where nitrate reduction predominately takes place in the mesophyll, and ammonia assimilation occurs in the bundle sheath (Figure S9). To our knowledge, the strong expression of genes associated with nitrate reduction has not previously been reported in any system. Furthermore, it has implications for our understanding of transitions associated with the evolution of the C_2 and C_4 pathways, as modelling revealed that early changes to C_2 metabolism should induce an imbalance in nitrogen metabolism between bundle sheath and mesophyll cells (Mallmann *et al.*, 2014). It has been proposed that this imbalance could be

© 2021 The Authors.

counteracted by upregulating genes associated with C_4 photosynthesis (Mallmann *et al.*, 2014). Indeed, transcripts for key genes of nitrate assimilation are consistently more strongly expressed in mesophyll cells of C_4 plants (Figure S9; Aubry *et al.*, 2014a; Chang *et al.*, 2012; Emms *et al.*, 2016; John *et al.*, 2014; Rao *et al.*, 2016). The consequences and the drivers of gene expression associated with nitrate reduction being focused on the bundle sheath in rice are unknown. However, this colocalization of sulphur and nitrate assimilation in the rice bundle sheath was associated with the preferential expression of specific ferredoxins that have previously only been implicated in differences in electron transfer reactions, allowing CO_2 or mineral nutrient reduction in shoots and roots (Yonekura-Sakakibara *et al.*, 2000).

The ancestors of rice and Arabidopsis diverged approximately 140 million years ago (Chaw *et al.*, 2004). Despite this timescale, significant overlap in cell-specific gene expression in bundle sheath and mesophyll cells from these two C_3 species was detected. We conclude that the two cell types have retained specific roles over this extended time. Less overlap was found in bundle sheath cells, suggesting that the role of the bundle sheath has evolved faster than the mesophyll. However, despite these apparent changes in gene expression in the bundle sheath, some of the genes associated with water transport and jasmonic acid biosynthesis were preferentially localized to the bundle sheath cells in most species. In contrast to sulphur assimilation, we therefore propose that these processes represent ancestral functions derived from the last common ancestor of monocotyledons and dicotyledons. Moreover, a small number of transcription factors were strongly localized to bundle sheath cells in both species, and so we propose that these regulators underpin ancestral and conserved functions of bundle sheath cells in dicotyledons and monocotyledons.

EXPERIMENTAL PROCEDURES

Plant growth conditions and sample preparation

The temperate rice (*O. sativa* ssp. *japonica*) Kitaake was germinated and grown in a mixture of 1:1 topsoil and sand for 2 weeks in a controlled environment growth room. The temperature was set to 28°C day/25°C night with a photoperiod of 12 h of light and 12 h of dark. The relative humidity was 60% and the photon flux density was 300 μmol m^{-2} sec^{-1}. Sections (1 cm) from the middle of the fourth fully expanded leaves were sampled 4 h after dawn. Leaf tissue was fixed and embedded into Steedman's wax as described previously (Hua and Hibberd, 2019), with minor modifications. For example, rice leaves were fixed in 100% (v/v) acetone on ice for 4 h and before embedding tissue infiltrated with 100% Steedman's wax at 37°C overnight. For LCM and RNA extraction, paradermal sections of 7 μm were prepared with a microtome and mounted on polyethylene naphthalate (PEN) membrane slides. Prior to LCM, Steedman's wax was removed by incubating slides in 100% (v/v) acetone for 1 min. LCM was performed on an Arcturus LCM platform (ThermoFisher Scientific, https://www.thermofisher.com), with isolated cells being collected on CapSure™ Macro Caps (ThermoFisher Scientific) and RNA extracted using the PicoPure™ RNA Isolation Kit (ThermoFisher Scientific) with on-column DNase I treatment. RNA quality and concentration were analysed using an Agilent Bioanalyser RNA 6000 Pico assay (Agilent, https://www.agilent.com).

For RNA *in situ* hybridization, middle sections from the fourth fully expanded leaf were sampled and fixed overnight using FAA fixative: 50% (v/v) ethanol, 5% (v/v) acetic acid and 3.7% (v/v) formaldehyde. They were then dehydrated through an ethanol series of 50, 70, 85, 95 and 100% (v/v) and embedded into Steedman's wax, as described previously (Hua and Hibberd, 2019). Sections (8-μm thick) were obtained on a microtome and mounted onto Superfrost plus slides. Tissue pre-treatment, hybridization and colour development were performed as previously described (Jackson, 1992), except that Steedman's wax was removed by incubating slides in 100% ethanol for 5 min twice.

Library preparation, RNA sequencing and data processing

20 ng of bundle sheath, mesophyll and veinal RNA was used as input for the Quantseq 3′ mRNA-seq library preparation kit (Moll *et al.*, 2014) according to the manufacturer's instructions (Lexogen, https://www.lexogen.com). Libraries were sequenced using Nextseq 500 sequencer to produce single-ended 150-bp reads for each sample. The leading and tailing 10-bp of Quantseq reads were trimmed and reads with a quality score of <20 and shorter than 50 bp were removed using BBDUK (Bushnell, 2014). Transcript abundance was determined after the remaining reads were quantified using SALMON 0.8.2 (Patro *et al.*, 2017) against the rice cDNA reference (MSU v7, http://rice.plantbiology.msu.edu) with the '--noLengthCorrection' flag used to disable length correction (Corley *et al.*, 2019). Gene-level abundance (transcripts per million, TPM) and counts were summarized using TXIMPORT 1.10.1 (Soneson *et al.*, 2016) and gene-level counts were used for the downstream differential gene expression analysis with DESEQ2 1.22.2 (Love *et al.*, 2014) and EDGER 3.24.3 (Robinson *et al.*, 2010). Poorly expressed genes with a row sum of TPM < 1 in three samples were excluded, resulting in 15 168 genes, and a Benjamini–Hochberg corrected *P* value (DESEQ2) and false discovery rate (FDR) (EDGER) of <0.05 were used to define differentially expressed genes.

To quantify the abundance of transcript associated with ribosomes of Arabidopsis BS and total leaf samples, reads were obtained from the Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra/), Bioproject accession PRJEB5030, and quantified using SALMON 0.8.2 against the Arabidopsis TAIR10 reference (Berardini *et al.*, 2015), using default parameters. Gene-level abundance (TPM) was summarized using TXIMPORT 1.10.1 (Soneson *et al.*, 2016). Differentially expressed genes were defined according to Aubry *et al.* (2014b).

Gene clustering, over-representation analysis and gene expression visualization

Consensus differentially expressed genes, identified using DESEQ2 and EDGER in each pairwise comparison, were used for gene expression clustering and functional enrichment analysis. Gene expression clustering was performed using the *K*-means method (Hartigan and Wong, 1979) with log_2-transformed quantile-normalized TPM values. For functional enrichment, the rice proteome (MSU v7) was first annotated by MAPMAN categories using MERCATOR 4 (Schwacke *et al.*, 2019), with over-representation analysis of MAPMAN categories performed using Fisher's exact test and the expressed 15 168 genes as a background, with FDR < 0.1. Rice transcription factor families were annotated according to PlantTFDB v5.0 (Jin *et al.*, 2017; http://planttfdb.gao-lab.org/),

© 2021 The Authors.
The Plant Journal published by Society for Experimental Biology and John Wiley & Sons Ltd,

using all expressed transcription factors as the background, and an over-representation test was performed using Fisher's exact test with a cut-off FDR of <0.1. Significant MAPMAN categories and transcription factor families were plotted using GGPLOT2 (Wickham, 2016). Heatmaps of gene expression clusters and pathways were plotted using COMPLEXHEATMAP (Gu *et al.*, 2016). Transcript abundance were presented in box plots of TPM values using the default settings of 'hinge' and 'whisker' in GGPLOT2.

Motif enrichment analysis

Leaf DNase I hypersensitive sites (DHSs) (Zhang *et al.*, 2012) within 2000 bp of the genes were extracted and used as inputs for the AME tool (http://meme-suite.org/tools/ame; Bailey *et al.*, 2015; McLeay and Bailey, 2010), using default settings and a custom background of all DHS regions within 2000 bp of the gene loci, as well as for the FIMO command line tool (Grant *et al.*, 2011), using default settings. The meme format Jaspar Plant Non-Redundant Motif Database (Fornes *et al.*, 2020; http://jaspar2018.genereg.net/downloads/) was used for both methods. FIMO scanning identified significant matches to known motifs and the frequencies of these sites were statistically tested for enrichment or depletion using permutation testing with the REGIONER package in R (Gel *et al.*, 2016). For permutation testing, all DHS regions within 2000 bp of the genes were used as a background for random subsampling and observed frequencies were statistically tested against the random subsampling distributions. As the Jaspar motif database contains many motifs with high similarity, predicted to be bound by closely related transcription factors, we grouped similar motifs together that showed a higher than 70% overlap in predicted target sites within the DHS background sequences, representing motifs that were deemed highly redundant. The most strongly enriched individual motif from each such group was used as a representative value for the group and plotted using GGPLOT2 (Wickham, 2016). In order to predict which *O. sativa* transcription factors might bind predicted motifs, we mapped rice transcription factors to their best match motif through protein homology with the Jaspar Non-Redundant Motif database proteins (Fornes *et al.*, 2020). The best-scoring BLASTP match was used with a bit-score of >100, to avoid spurious assignments to motifs from other transcription factor families. For all of the transcription factors found in the six clusters, those that mapped to an enriched motif from any of the six clusters were plotted using the single motif enrichment scores and rice transcription factors were annotated according to the funRiceGenes database (Yao *et al.*, 2018).

Defining orthogroups in rice and Arabidopsis

Protein sequences of primary transcripts of *A. thaliana* (TAIR10), *Brachypodium distachyon* (v3.2), *G. gynandra* (Aubry *et al.*, 2014a), *Marchantia polymorpha* (v3.1), *O. sativa* (v7.0), *P. virgatum* (v1.1), *Setaria italica* (v2.1), *Sorghum bicolor* (v1.4) and *Z. mays* (Ensembl-18), obtained from Phytozome v12.1 (Goodstein *et al.*, 2012) were clustered into orthogroups using ORTHOFINDER 2.4.1 (Emms and Kelly, 2019) with default parameters, and orthogroups that contain both rice and Arabidopsis genes were used for comparative analysis in the two species. Resolved gene trees of the orthogroups were visualized using the GGTREE 2.41 package in R (Yu, 2020).

Accession numbers

Raw sequencing data are deposited in the National Center for Biotechnology Information under BioProject ID PRJNA702624, BioSample ID SAMN17976370–SAMN17976383.

ACKNOWLEDGEMENTS

The work was funded by a C_4 Rice project grant from the Bill and Melinda Gates Foundation to the University of Oxford (2015–2019), European Research Council Grant 694733 Revolution, and BBSRC grants BBP0031171 and BBL014130 to JMH. Research in SK's lab is funded by the Deutsche Forschungsgemeinschaft (DFG) under Germany's Excellence Strategy EXC 2048/1 project 390686111.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest associated with this work.

DATA AVAILABILITY STATEMENT

Code associated with this article and the underlying data required to generate plots are available in the GitHub repository: https://github.com/hibberd-lab/Hua_et_al_2021_Kitaake_LCM_RNAseq. All other data are available on request.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Transcript abundance of genes previously reported to be associated with mesophyll, bundle sheath or veinal cells.

Figure S2. MAPMAN categories and metabolic overview of pairwise comparisons between bundle sheath (BS) and vein (V) and between mesophyll (M) and vein (V).

Figure S3. Enriched transporter families in the six gene expression clusters.

Figure S4. Transcript abundance of genes associated with nitrogen assimilation.

Figure S5. Transcript abundance of genes associated with the photosynthetic electron transport chain (a, b), Calvin–Benson–Bassham cycle (c), starch biosynthesis (d) and photorespiration (e).

Figure S6. Transcription factors associated with different clusters.

Figure S7. Transcript abundance of orthogroups shared by bundle sheath cells in rice and Arabidopsis associated with aquaporins, sulphate transport and assimilation, jasmonic acid biosynthesis, and transcription factors.

Figure S8. Comparison of bundle sheath preferentially expressed genes among different studies.

Figure S9. Transcript abundance of aquaporins (*PIP1* and *PIP2*), ATP sulfurylase (*ATPS*), APS reductase (*APR*), sulfite reductase (*SIR*), 13-lipoxygenase (*LOX*), allene oxidase cyclase (*AOC*), oxophytodienoate reductase (*OPR*), nitrate reductase (*NIA*) and nitrite reductase (*NIR*) in bundle sheath and mesophyll cells of *A. thaliana* (Aubry *et al.*, 2014b), *G. gynandra* (Aubry *et al.*, 2014a), *O. sativa* (this study), *P. virgartum* (Rao *et al.*, 2016), *Setaria italica* (John *et al.*, 2014), *Sorghum bicolor* (Emms *et al.*, 2016) and *Z. mays* (Chang *et al.*, 2012).

Table S1. RNA sequencing statistics.

Table S2. Summary of differential gene expression analysis using DESEQ2 and EDGER.

Table S3. Pairwise comparison and gene expression clusters.

Table S4. Motif enrichment analysis of the six gene expression clusters.

Table S5. Bundle sheath and mesophyll differentially expressed orthologous genes in rice and Arabidopsis.

© 2021 The Authors.

REFERENCES

Attia, Z., Dalal, A. & Moshelion, M. (2020) Vascular bundle sheath and mesophyll cells modulate leaf water balance in response to chitin. *The Plant Journal*, **101**, 1368–1377.

Aubry, S., Kelly, S., Kümpers, B.M.C.C., Smith-Unna, R.D. & Hibberd, J.M. (2014a) Deep evolutionary comparison of gene expression identifies parallel recruitment of trans-factors in two independent origins of C_4 photosynthesis. *PLoS Genetics*, **10**, e1004365.

Aubry, S., Smith-Unna, R.D., Boursnell, C.M., Kopriva, S. & Hibberd, J.M. (2014b) Transcript residency on ribosomes reveals a key role for the *Arabidopsis thaliana* bundle sheath in sulfur and glucosinolate metabolism. *The Plant Journal*, **78**, 659–673.

Bailey, T.L., Johnson, J., Grant, C.E. & Noble, W.S. (2015) The MEME suite. *Nucleic Acids Research*, **43**, 39–49.

Berardini, T.Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E. *et al.* (2015) The *Arabidopsis* information resource: Making and mining the "gold standard" annotated reference plant genome. *Genesis*, **53**, 474–485.

Berry, J.A. (2012) There ought to be an equation for that. *Annual Review of Plant Biology*, **63**, 1–17.

Bezrutczyk, M., Zöllner, N.R., Kruse, C.P.S., Hartwig, T., Lautwein, T., Köhrer, K. *et al.* (2021) Evidence for phloem loading via the abaxial bundle sheath cells in maize leaves. *The Plant Cell*. https://doi.org/10.1093/plcell/koaa055.

Birnbaum, K., Shasha, D.E., Wang, J.Y., Jung, J.W., Lambert, G.M., Galbraith, D.W. *et al.* (2003) A gene expression map of the *Arabidopsis* root. *Science*, **302**, 1956–1960.

Brady, S.M., Orlando, D.A., Lee, J.Y., Wang, J.Y., Koch, J., Dinneny, J.R. *et al.* (2007) A high-resolution root spatiotemporal map reveals dominant expression patterns. *Science*, **318**, 801–806.

Brown, N.J., Palmer, B.G., Stanley, S., Hajaji, H., Janacek, S.H., Astley, H.M. *et al.* (2010) C_4 acid decarboxylases required for C_4 photosynthesis are active in the mid-vein of the C_3 species *Arabidopsis thaliana*, and are important in sugar and amino acid metabolism. *The Plant Journal*, **61**, 122–133.

Burgener, M., Suter, M., Jones, S. & Brunold, C. (1998) Cyst(e)ine is the transport metabolite of assimilated sulfur from bundle-sheath to mesophyll cells in maize leaves. *Plant Physiology*, **116**, 1315–1322.

Burnell, J.N. (1984) Sulfate assimilation in C_4 plants. *Plant Physiology*, **75**, 873–875.

Bushnell, B. (2014) *BBMap: A fast, accurate, splice-aware aligner*. LBNL Report #: LBNL-7065E. Lawrence Berkeley National Laboratory. Available from: https://escholarship.org/uc/item/1h3515gn [Accessed 5th January 2019].

Chang, Y.M., Liu, W.Y., Shih, A.C.C., Shen, M.N., Lu, C.H., Lu, M.Y.J. *et al.* (2012) Characterizing regulatory and functional differentiation between maize mesophyll and bundle sheath cells by transcriptomic analysis. *Plant Physiology*, **160**, 165–177.

Chávez-Bárcenas, A.T., Valdez-Alarcón, J.J., Martínez-Trujillo, M., Chen, L., Xoconostle-Cázares, B., Lucas, W.J. *et al.* (2000) Tissue-specific and developmental pattern of expression of the rice *sps1* gene. *Plant Physiology*, **124**, 641–654.

Chaw, S.M., Chang, C.C., Chen, H.L. & Li, W.H. (2004) Dating the monocot-dicot divergence and the origin of core eudicots using whole chloroplast genomes. *Journal of Molecular Evolution*, **58**, 424–441.

Chen, T., Wu, H., Wu, J., Fan, X., Li, X. & Lin, Y. (2017) Absence of *OsβCA1* causes a CO_2 deficit and affects leaf photosynthesis and the stomatal response to CO_2 in rice. *The Plant Journal*, **90**, 344–357.

Cho, J.I., Burla, B., Lee, D.W., Ryoo, N., Hong, S.K., Kim, H.B. *et al.* (2010) Expression analysis and functional characterization of the monosaccharide transporters, *OsTMTs*, involving vacuolar sugar transport in rice. *New Phytologist*, **186**, 657–668.

Corley, S.M., Troy, N.M., Bosco, A. & Wilkins, M.R. (2019) QuantSeq 3′ Sequencing combined with Salmon provides a fast, reliable approach for high throughput RNA expression analysis. *Scientific Reports*, **9**, 1–15.

d Moore, B.R., Ku, M.S. & Edwards, G.E. (1984) Isolation of leaf bundle sheath protoplasts from C_4 dicot species and intracellular localization of selected enzymes. *Plant Science Letters*, **35**, 127–138.

Dickinson, P.J., Kneřová, J., Szecówka, M., Stevenson, S.R., Burgess, S.J., Mulvey, H. *et al.* (2020) A bipartite transcription factor module controlling expression in the bundle sheath of *Arabidopsis thaliana*. *Nature Plants*, **6**, 1468–1479.

Doyama, N., Kajiwara, H. & Ida, S. (1998) Cloning and expression of a ferredoxin gene in rice roots in response to nitrate and ammonium. *Plant Science*, **137**, 53–62.

Edwards, G.E. & Black, C.C. (1971) Isolation of mesophyll cells and bundle sheath cells from *Digitaria sanguinalis* (L.) Scop. leaves and a scanning microscopy study of the internal leaf cell morphology. *Plant Physiology*, **47**, 149–156.

Emms, D.M., Covshoff, S., Hibberd, J.M. & Kelly, S. (2016) Independent and parallel evolution of new genes by gene duplication in two origins of C_4 photosynthesis provides new insight into the mechanism of phloem loading in C_4 species. *Molecular Biology and Evolution*, **33**, 1796–1806.

Emms, D.M. & Kelly, S. (2019) OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology*, **20**, 238.

Fetter, K., Van Wilder, V., Moshelion, M. & Chaumont, F. (2004) Interactions between plasma membrane aquaporins modulate their water channel activity. *The Plant Cell*, **16**, 215–228.

Fornes, O., Castro-Mondragon, J.A., Khan, A., Van der Lee, R., Zhang, X., Richmond, P.A. *et al.* (2020) JASPAR 2020: update of the open-Access database of transcription factor binding profiles. *Nucleic Acids Research*, **48**, D87–D92.s

Foyer, C.H. & Noctor, G. (2011) Ascorbate and glutathione: the heart of the redox hub. *Plant Physiology*, **155**, 2–18.

Fryer, M.J., Oxborough, K., Mullineaux, P.M. & Baker, N.R. (2002) Imaging of photo-oxidative stress responses in leaves. *Journal of Experimental Botany*, **53**, 1249–1254.

Furbank, R.T. (2016) Walking the C_4 pathway: past, present, and future. *Journal of Experimental Botany*, **67**, 4057–4066.

Gel, B., Díez-Villanueva, A., Serra, E., Buschbeck, M., Peinado, M.A. & Malinverni, R. (2016) RegioneR: an R/Bioconductor package for the association analysis of genomic regions based on permutation tests. *Bioinformatics*, **32**, 289–291.

Gerwick, B.C., Ku, S.B. & Black, C.C. (1980) Initiation of sulfate activation: a variation in C_4 photosynthesis plants. *Science*, **209**, 513–515.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J. *et al.* (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research*, **40**, D1178.

Grant, C.E., Bailey, T.L. & Noble, W.S. (2011) FIMO: scanning for occurrences of a given motif. *Bioinformatics*, **27**, 1017–1018.

Griffiths, H., Weller, G., Toy, L.F.M. & Dennis, R.J. (2013) You're so vein: bundle sheath physiology, phylogeny and evolution in C_3 and C_4 plants. *Plant, Cell and Environment*, **36**, 249–261.

Grunwald, Y., Wigoda, N., Sade, N., Yaaran, A., Torne, T., Chaka Gosa, S. *et al.* (2021) Arabidopsis leaf hydraulic conductance is regulated by xylem-sap pH, controlled, in turn, by a P-type H^+-ATPase of vascular bundle sheath cells. *The Plant Journal*, **106**, 301–313.

Gu, Z., Eils, R. & Schlesner, M. (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, **32**, 2847–2849.

Haberlandt, G. (1884). *Physiologische Pflanzenanatomie*. Leipzig: Wilhelm Engelmann.

Hartigan, J.A. & Wong, M.A. (1979) A K-means clustering algorithm. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **28**, 100.

Hatch, M.D. (1987) C_4 photosynthesis: a unique elend of modified biochemistry, anatomy and ultrastructure. *Biochimica et Biophysica Acta (BBA)-Reviews on Bioenergetics*, **895**, 81–106.

Hatch, M.D. & Burnell, J.N. (1990) Carbonic anhydrase activity in leaves and its role in the first step of C_4 photosynthesis. *Plant Physiology*, **93**, 825–828.

He, L., Li, M., Qiu, Z., Chen, D., Zhang, G., Wang, X. *et al.* (2020) Primary leaf-type ferredoxin 1 participates in photosynthetic electron transport and carbon assimilation in rice. *The Plant Journal*, **104**, 44–58.

Hernández, L.E., Sobrino-Plata, J., Belén Montero-Palmero, M., Carrasco-Gil, S., Flores-Cáceres, L., Ortega-Villasante, C. *et al.* (2015) Contribution of glutathione to the control of cellular redox homeostasis under toxic metal and metalloid stress. *Journal of Experimental Botany*, **66**, 2901–2911.

Hibberd, J.M. & Quick, W.P. (2002) Characteristics of C_4 photosynthesis in stems and petioles of C_3 flowering plants. *Nature*, **415**, 451–454.

Hibberd, J.M., Sheehy, J.E. & Langdale, J.A. (2008) Using C_4 photosynthesis to increase the yield of rice—rationale and feasibility. *Current Opinion in Plant Biology*, **11**, 228–231.

© 2021 The Authors.
The Plant Journal published by Society for Experimental Biology and John Wiley & Sons Ltd,

Hua, L. & Hibberd, J.M. (2019) An optimized protocol for isolation of high-quality RNA through laser capture microdissection of leaf material. *Plant Direct*, **3**, e00156.

Ibraheem, O., Botha, C.E., Bradley, G., Dealtry, G. & Roux, S. (2013) Rice sucrose transporter1 (OsSUT1) up-regulation in xylem parenchyma is caused by aphid feeding on rice leaf blade vascular bundles. *Plant Biology*, **16**, 783–791.

Jackson, D. (1992) In situ hybridization in plants. In: *Molecular Plant Pathology: A Practical Approach*. Oxford University Press, pp. 163–174.

Janacek, S.H., Trenkamp, S., Palmer, B., Brown, N.J., Parsley, K., Stanley, S. *et al.* (2009) Photosynthesis in cells around veins of the C_3 plant *Arabidopsis thaliana* is important for both the shikimate pathway and leaf senescence as well as contributing to plant fitness. *The Plant Journal*, **59**, 329–343.

Jiao, Y., Tausta, S.L., Gandotra, N., Sun, N., Liu, T., Clay, N.K. *et al.* (2009) A transcriptome atlas of rice cell types uncovers cellular, functional and developmental hierarchies. *Nature Genetics*, **41**, 258–263.

Jin, J., Tian, F., Yang, D.C., Meng, Y.Q., Kong, L., Luo, J. *et al.* (2017) PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Research*, **45**, D1040–D1045.

John, C.R., Smith-Unna, R.D., Woodfield, H., Covshoff, S. & Hibberd, J.M. (2014) Evolutionary convergence of cell-specific gene expression in independent lineages of C_4 grasses. *Plant Physiology*, **165**, 62–75.

Jurić, I., González-Pérez, V., Hibberd, J.M., Edwards, G. & Burroughs, N.J. (2016) Size matters for single-cell C_4 photosynthesis in *Bienertia*. *Journal of Experimental Botany*, **68**, 255–267.

Kanai, R. & Edwards, G.E. (1973) Separation of mesophyll protoplasts and bundle sheath cells from maize leaves for photosynthetic studies. *Plant Physiology*, **51**, 1133–1137.

Kim, J.-Y., Symeonidi, E., Pang, T.Y., Denyer, T., Weidauer, D., Bezrutczyk, M. *et al.* (2021) Distinct identities of leaf phloem cells revealed by single cell transcriptomics. *The Plant Cell*. https://doi.org/10.1093/plcell/koaa060.

Kinsman, E.A. & Pyke, K.A. (1998) Bundle sheath cells and cell-specific plastid development in *Arabidopsis* leaves. *Development*, **125**, 1815–1822.

Kopriva, S. & Koprivova, A. (2005) Sulfate assimilation and glutathione synthesis in C_4 plants. *Photosynthesis Research*, **86**, 363–372.

Koprivova, A., Melzer, M., Von Ballmoos, P., Mandel, T., Brunold, C. & Kopriva, S. (2001) Assimilatory sulfate reduction in C_3, C_3–C_4, and C_4 species of *Flaveria*. *Plant Physiology*, **127**, 543–550.

Koroleva, O.A., Gibson, T.M., Cramer, R. & Stain, C. (2010) Glucosinolate-accumulating S-cells in *Arabidopsis* leaves and flower stalks undergo programmed cell death at early stages of differentiation. *The Plant Journal*, **64**, 456–469.

Langdale, J.A. (2011) C_4 cycles: past, present, and future research on C_4 photosynthesis. *The Plant Cell*, **23**, 3879–3892.

Leegood, R.C. (2008) Roles of the bundle sheath cells in leaves of C_3 plants. *Journal of Experimental Botany*, **59**, 1663–1673.

Li, G.W., Zhang, M.H., Cai, W.M., Sun, W.N. & Su, W.A. (2008) Characterization of OsPIP2;7, a water channel protein in rice. *Plant and Cell Physiology*, **49**, 1851–1858.

Love, M.I., Huber, W. & Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**, 550.

Lunn, J.E. & Furbank, R.T. (1997) Localisation of sucrose-phosphate synthase and starch in leaves of C_4 plants. *Planta*, **202**, 106–111.

Mallmann, J., Heckmann, D., Bräutigam, A., Lercher, M.J., Weber, A.P.M., Westhoff, P. *et al.* (2014) The role of photorespiration during the evolution of C_4 photosynthesis in the genus *Flaveria*. *Elife*, **3**, e02478.

Maruyama-Nakashita, A., Nakamura, Y., Tohge, T., Saito, K. & Takahashi, H. (2006) *Arabidopsis* SLIM1 is a central transcriptional regulator of plant sulfur response and metabolism. *The Plant Cell*, **18**, 3235–3251.

Masumoto, C., Miyazawa, S.-I.-I., Ohkawa, H., Fukuda, T., Taniguchi, Y., Murayama, S. *et al.* (2010) Phospho*enol*pyruvate carboxylase intrinsically located in the chloroplast of rice plays a crucial role in ammonium assimilation. *Proceedings of the National Academy of Sciences of United States of America*, **107**, 5226–5231.

Maughan, S.C., Pasternak, M., Cairns, N., Kiddle, G., Brach, T., Jarvis, R. *et al.* (2010) Plant homologs of the *Plasmodium falciparum* chloroquine-resistance transporter, *Pf*CRT, are required for glutathione homeostasis and stress responses. *Proceedings of the National Academy of Sciences of United States of America*, **107**, 2331–2336.

McLeay, R.C. & Bailey, T.L. (2010) Motif enrichment analysis: a unified framework and an evaluation on ChIP data. *BMC Bioinformatics*, **11**, 165.

Miyake, H. (2016) Starch accumulation in the bundle sheaths of C_3 plants: a possible pre-condition for C_4 photosynthesis. *Plant and Cell Physiology*, **57**, 890–896.

Moll, P., Ante, M., Seitz, A. & Reda, T. (2014) QuantSeq 3′ mRNA sequencing for RNA quantification. *Nature Methods*, **11**, i–iii.

Mustroph, A., Zanetti, M.E., Jang, C.J.H.H., Holtan, H.E., Repetti, P.P., Galbraith, D.W. *et al.* (2009) Profiling translatomes of discrete cell populations resolves altered cellular priorities during hypoxia in *Arabidopsis*. *Proceedings of the National Academy of Sciences of United States of America*, **106**, 18843–18848.

Nomura, M., Higuchi, T., Ishida, Y., Ohta, S., Komari, T., Imaizumi, N. *et al.* (2005) Differential expression pattern of C_4 bundle sheath expression genes in rice, a C_3 plant. *Plant and Cell Physiology*, **46**, 754–761.

Passera, C. & Ghisi, R. (1982) ATP sulphurylase and O-acetylserine sulphydrylase in isolated mesophyll protoplasts and bundle sheath strands of S-deprived maize leaves. *Journal of Experimental Botany*, **33**, 432–438.

Pasternak, M., Lim, B., Wirtz, M., Hell, R., Cobbett, C.S. & Meyer, A.J. (2007) Restricting glutathione biosynthesis to the cytosol is sufficient for normal plant development. *The Plant Journal*, **53**, 999–1012.

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A. & Kingsford, C. (2017) Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, **14**, 417–419.

Rao, X., Lu, N., Li, G., Nakashima, J., Tang, Y. & Dixon, R.A. (2016) Comparative cell-specific transcriptomics reveals differentiation of C_4 photosynthesis pathways in switchgrass and other C_4 lineages. *Journal of Experimental Botany*, **67**, 1649–1662.

Robinson, M.D., Mccarthy, D.J. & Smyth, G.K. (2010) edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139–140.

Sade, N., Shatil-Cohen, A., Attia, Z., Maurel, C., Boursiac, Y., Kelly, G. *et al.* (2014) The role of plasma membrane aquaporins in regulating the bundle sheath-mesophyll continuum and leaf hydraulics. *Plant Physiology*, **166**, 1609–1620.

Sage, R.F. (2001) Environmental and evolutionary preconditions for the origin and diversification of the C_4 photosynthetic syndrome. *Plant Biology*, **3**, 202–213.

Sage, R.F. (2004) The evolution of C_4 photosynthesis. *New Phytologist*, **161**, 341–370.

Sage, R.F., Sage, T.L. & Kocacinar, F. (2012) Photorespiration and the evolution of C_4 photosynthesis. *Annual Review of Plant Biology*, **63**, 19–47.

Sage, T.L. & Sage, R.F. (2009) The functional anatomy of rice leaves: implications for refixation of photorespiratory CO_2 and efforts to engineer C_4 photosynthesis into rice. *Plant and Cell Physiology*, **50**, 756–772.

Sakurai, J., Ishikawa, F., Yamaguchi, T., Uemura, M. & Maeshima, M. (2005) Identification of 33 rice aquaporin genes and analysis of their expression and function. *Plant and Cell Physiology*, **46**, 1568–1577.

Sawers, R.J.H., Liu, P., Anufrikova, K., Gene, J.T.G. & Brutnell, T.P. (2007) A multi-treatment experimental system to examine photosynthetic differentiation in the maize leaf. *BMC Genomics*, **8**, 12.

Schmutz, D. & Brunold, C. (1984) Intercellular localization of assimilatory sulfate reduction in leaves of *Zea mays* and *Triticum aestivum*. *Plant Physiology*, **74**, 866–870.

Schwacke, R., Ponce-Soto, G.Y., Krause, K., Bolger, A.M., Arsova, B., Hallab, A. *et al.* (2019) MapMan4: a refined protein classification and annotation framework applicable to multi-omics data analysis. *Molecular Plant*, **12**, 879–892.

Scofield, G.N., Hirose, T., Aoki, N. & Furbank, R.T. (2007) Involvement of the sucrose transporter, OsSUT1, in the long-distance pathway for assimilate transport in rice. *Journal of Experimental Botany*, **58**, 3155–3169.

Shatil-Cohen, A., Attia, Z. & Moshelion, M. (2011) Bundle-sheath cell regulation of xylem-mesophyll water transport via aquaporins under drought stress: a target of xylem-borne ABA? *The Plant Journal*, **67**, 72–80.

Shen, W., Ye, L., Ma, J., Yuan, Z., Zheng, B., Chuangen, L.V. *et al.* (2016) The existence of C_4-bundle-sheath-like photosynthesis in the mid-vein of C_3 rice. *Rice*, **9**, 20.

Soneson, C., Love, M.I. & Robinson, M.D. (2016) Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research*, **4**, 1521.

© 2021 The Authors.

Taylor, L., Nunes-Nesi, A., Parsley, K., Leiss, A., Leach, G., Coates, S. *et al.* (2010) Cytosolic pyruvate, orthophosphate dikinase functions in nitrogen remobilization during leaf senescence and limits individual seed growth and nitrogen content. *The Plant Journal*, **62**, 641–652.

Thimm, O., Bläsing, O., Gibon, Y., Nagel, A., Meyer, S., Krüger, P. *et al.* (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *The Plant Journal*, **37**, 914–939.

Tipple, B.J. & Pagani, M. (2007) The early origins of terrestrial C_4 photosynthesis. *Annual Review of Earth and Planetary Sciences*, **35**, 435–461.

von Caemmerer, S., Edwards, G.E., Koteyeva, N. & Cousins, A.B. (2014) Single cell C_4 photosynthesis in aquatic and terrestrial plants: a gas exchange perspective. *Aquatic Botany*, **118**, 71–80.

von Caemmerer, S., Quick, W.P. & Furbank, R.T. (2012) The development of C_4 rice: current progress and future challenges. *Science*, **336**, 1671–1672.

Voznesenskaya, E.V., Franceschi, V.R., Kiirats, O., Freitag, H. & Edwards, G.E. (2001) Kranz anatomy is not essential for terrestrial C_4 plant photosynthesis. *Nature*, **414**, 543–546.

Wang, P., Khoshravesh, R., Karki, S., Tapia, R., Balahadia, C.P., Bandyopadhyay, A. *et al.* (2017) Re-creation of a key step in the evolutionary switch from C_3 to C_4 leaf anatomy. *Current Biology*, **27**, 3278–3287.e6.

Waters, M.T., Moylan, E.C. & Langdale, J.A. (2008) GLK transcription factors regulate chloroplast development in a cell-autonomous manner. *The Plant Journal*, **56**, 432–444.

Waters, M.T., Wang, P., Korkaric, M., Capper, R.G., Saunders, N.J. & Langdale, J.A. (2009) GLK transcription factors coordinate expression of the photosynthetic apparatus in *Arabidopsis*. *The Plant Cell*, **21**, 1109–1128.

Wickham, H. (2016). *ggplot2: elegant graphics for data analysis*. New York: Springer- Verlag.

Wigoda, N., Pasmanik-Chor, M., Yang, T., Yu, L., Moshelion, M. & Moran, N. (2017) Differential gene expression and transport functionality in the bundle sheath versus mesophyll-a potential role in leaf mineral homeostasis. *Journal of Experimental Botany*, **68**, 3179–3190.

Williams, M.L., Farrar, J.F. & Pollock, C.J. (1989) Cell specialization within the parenchymatous bundle sheath of barley. *Plant, Cell and Environment*, **12**, 909–918.

Yang, J., Gao, M., Hu, H., Ding, X., Lin, H., Wang, L. *et al.* (2016) OsCLT1, a CRT-like transporter 1, is required for glutathione homeostasis and arsenic tolerance in rice. *New Phytologist*, **211**, 658–670.

Yao, W., Li, G., Yu, Y. & Ouyang, Y. (2018) funRiceGenes dataset for comprehensive understanding and application of rice functional genes. *Gigascience*, **7**, 1–9.

Yonekura-Sakakibara, K., Onda, Y., Ashikari, T., Tanaka, Y., Kusumi, T. & Hase, T. (2000) Analysis of reductant supply systems for ferredoxin-dependent sulfite reductase in photosynthetic and nonphotosynthetic organs of maize. *Plant Physiology*, **122**, 887–894.

Yu, G. (2020) Using ggtree to visualize data on tree-like structures. *Current Protocols in Bioinformatics*, **69**(1), e96.

Zelazny, E., Borst, J.W., Muylaert, M., Batoko, H., Hemminga, M.A. & Chaumont, F. (2007) FRET imaging in living maize cells reveals that plasma membrane aquaporins interact to regulate their subcellular localization. *Proceedings of the National Academy of Sciences of United States of America*, **104**, 12359–12364.

Zhang, W., Wu, Y., Schnable, J.C., Zeng, Z., Freeling, M., Crawford, G.E. *et al.* (2012) High-resolution mapping of open chromatin in the rice genome. *Genome Research*, **22**, 151–162.

© 2021 The Authors.
The Plant Journal published by Society for Experimental Biology and John Wiley & Sons Ltd,